

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 20, 2004, 16:41:26 ; Search time 33.8015 Seconds
(without alignments)
42.263 Million cell updates/sec

Title: US-10-089-549-11

Perfect score: 44

Sequence: 1 HVSALEGNLL 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_19Jun03.*
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	44	100.0	9	23	Alpha-isomaltosylg
2	44	100.0	1251	23	Alpha-isomaltosylg
3	44	100.0	1286	23	Alpha-isomaltosylg
4	40	90.9	8	23	Alpha-isomaltosylg
5	37	84.1	330	22	C glutamicum prote
6	37	84.1	380	22	Corynebacterium g
7	35	79.5	162	21	Zea mays protein f
8	35	79.5	196	21	Zea mays protein f
9	35	79.5	3931	24	Human protein NOV9

10	34	77.3	109	22	AAG90316	C glutamicum prote
11	34	77.3	526	24	ABJ25590	Aspergillus fumiga
12	33	75.0	9	23	AB84268	B. globisporus deh
13	33	75.0	9	23	ABG30540	Alpha-isomaltosylg
14	33	75.0	9	23	AAO17330	Alpha-isomaltosylg
15	33	75.0	188	22	AAU57279	Propionibacterium
16	33	75.0	1249	23	ABG30537	Alpha-isomaltosylg
17	33	75.0	1284	23	ABG30563	Alpha-isomaltosylg
18	33	75.0	1284	24	ABP70652	Bacillus globispor
19	33	75.0	1284	24	ABP57724	protein #1 related
20	32	72.7	130	23	ABP30687	Streptococcus poly
21	32	72.7	147	22	ABN96275	Putative diadenosi
22	32	72.7	243	22	ABG11899	Novel human diagno
23	32	72.7	292	22	ABG11900	Novel human diagno
24	32	72.7	312	22	ABG63330	Drosophila melancog
25	32	72.7	312	22	ABG63378	Drosophila atonal
26	32	72.7	550	22	ABN71906	Drosophila melancog
27	32	72.7	1578	19	AAW69361	Tetradotoxin-sensi
28	32	72.7	1980	21	AAJ23563	Human sodium chann
29	32	72.7	1980	23	AAJ14927	Human sodium chann
30	32	72.7	1988	19	AAW69362	Tetradotoxin-sensi
31	31	70.5	51	23	ABP01357	Human ORFX protein
32	31	70.5	58	23	ABP01321	Human ORFX protein
33	31	70.5	66	22	AAW82549	Human immune/haema
34	31	70.5	104	22	ABG49590	Human liver peptid
35	31	70.5	104	22	ABB29583	Peptide #2234 enco
36	31	70.5	104	22	ABB34762	Peptide #2268 enco
37	31	70.5	104	22	ABB20177	Protein #2176 enco
38	31	70.5	104	22	AAW55564	Human brain expres
39	31	70.5	104	22	AAW67948	Human bone marrow
40	31	70.5	104	22	AAW15766	Peptide #2200 enco
41	31	70.5	104	22	AAW28275	Peptide #2312 enco
42	31	70.5	104	22	AAW03500	Peptide #2182 enco
43	31	70.5	104	23	ABG37483	Human peptide enco
44	31	70.5	106	23	ABP35084	Human transport pr
45	31	70.5	130	22	ABG08777	Novel human diagno

ALIGNMENTS

RESULT 1
AAO17340
ID AAO17340 standard; peptide; 9 AA.
XX
AC AAO17340;
XX
DT 08-JUL-2002 (first entry)
XX
DE Alpha-isomaltosylglucosaccharide synthase related peptide #11.
XX
KW Alpha-isomaltosylglucosaccharide synthase; enzyme; sweeter; food;
KW drink; cosmetics; pharmaceutical; cyclic tetrasaccharide.
XX
OS Bacillus globisporus.
XX
PN WO200210361-A1.
XX
PD 07-FEB-2002.
XX
PF 25-JUL-2001; 2001WO-JP06412.
XX
PR 01-AUG-2000; 2000JP-0233364.
PR 02-AUG-2000; 2000JP-0234937.
(HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
XX
PI Kubota M, Tsusaki K, Higashiyama T, Fukuda S, Miyake T;
XX
DR WPI; 2002-315302/35.
XX
PT Alpha-isomaltosylglucosaccharide synthase capable of transferring
PT Alpha-glucosyl from saccharide, useful in producing cyclic

PT tetrasaccharides and related carbohydrates industrially for application
PT in compositions e.g. as drugs -
XX
PS Claim 5; Page 203; 209pp; Japanese.
XX
CC The present invention relates to an alpha-isomaltosylglucosaccharide
CC synthase which is capable of transferring alpha-glucosyl from a
CC saccharide to form a specific saccharide which carries an alpha-1,6
CC glucosyl linkage at its non-reducing end and an alpha-1,4 glucosyl
CC linkage at ends other than the non-reducing end and has a degree of
CC glucose polymerization of at most 3, but without substantially elevating
CC the reducing ability. The synthase is useful in producing cyclic
CC tetrasaccharides and related carbohydrates industrially for application
CC in compositions e.g. as sweeteners, low calorie food materials, taste
CC improvers, flavour improvers, quality improvers, water-separation
CC inhibitors, stabilizers, excipients, binding agents and pulverisation
CC bases, especially in foods, drinks, cosmetics and pharmaceuticals. The
CC present sequence is a peptide described in the exemplification of the
CC invention.
XX
XX Sequence 9 AA;
SQ
Query Match 100.0%; Score 44; DB 23; Length 9;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 HVSAALGNLL 9
DB 1 HVSAALGNLL 9

RESULT 2
ABG30538
ID ABG30538 standard; Protein; 1251 AA.
XX
AC ABG30538;
XX
DT 07-OCT-2002 (first entry)
XX
DE Alpha-isomaltosylglucosaccharide synthase #2 mature protein.
XX
KW Alpha-isomaltosylglucosaccharide synthase; sugar production;
KW cyclic tetrasaccharide; cyclo(-6)-alpha-D-glucopyranosyl(1-3);
KW alpha-isomaltosyltransferase; viscosity modifier; sweetener;
KW heat stabiliser; hydrolysis stabiliser; colour stabiliser; foodstuff;
KW cosmetic; drug composition; enzyme.
XX
OS Unidentified.
XX
PN WO200255708-A1.
XX
PD 18-JUL-2002.
XX
PF 09-JAN-2002; 2002WO-JP00052.
XX
PR 12-JAN-2001; 2001JP-0005441.
XX
PA (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
XX
PI Kubota M, Maruta K, Yamamoto T, Fukuda S;
XX
DR WPI; 2002-520129/55.
DR N-PSDB; ABK88158.
XX
PT New alpha-isomaltosylglucosaccharide synthase of bacterial origin for
PT the production of cyclic tetrasaccharide gum -
XX
PS Claim 1; Page 105-109; 144pp; Japanese.
XX
CC The invention describes novel microbial polypeptides having
CC alpha-isomaltosylglucosaccharide synthase activity. The proteins are
CC useful for producing a sugar (I) having at least three glucose units
CC connected by alpha-1,4 bonds except for alpha-1,6 at the non-reducing
CC end. The invention also describes a method for the production of the
CC cyclic tetrasaccharide cyclo(-6)-alpha-D-glucopyranosyl(1-3)-alpha-D-
CC glucopyranosyl(1-6)-alpha-D-glucopyranosyl(1-3)-alpha-D-glucopyranosyl
CC (1-) (II) by treatment of (I) with alpha-isomaltosyltransferase.
CC (II) and similar sugars in crystalline or syrup form are used as
CC a viscosity modifier, sweetener, heat stabiliser, hydrolysis stabiliser
CC and colour stabiliser for foodstuffs, cosmetics and drug compositions.
CC This is the amino acid sequence of an alpha-isomaltosylglucosaccharide
CC synthase.

CC end. The invention also describes a method for the production of the
CC cyclic tetrasaccharide cyclo(-6)-alpha-D-glucopyranosyl(1-3)-alpha-D-
CC glucopyranosyl(1-6)-alpha-D-glucopyranosyl(1-3)-alpha-D-glucopyranosyl
CC (1-) (II) by treatment of (I) with alpha-isomaltosyltransferase.
CC (II) and similar sugars in crystalline or syrup form are used as
CC a viscosity modifier, sweetener, heat stabiliser, hydrolysis stabiliser
CC and colour stabiliser for foodstuffs, cosmetics and drug compositions.
CC This is the amino acid sequence of an alpha-isomaltosylglucosaccharide
CC synthase mature protein.
XX
SQ Sequence 1251 AA;
Query Match 100.0%; Score 44; DB 23; Length 1251;
Best Local Similarity 100.0%; Pred. No. 3.6;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 HVSAALGNLL 9
DB 1 HVSAALGNLL 9

RESULT 3
ABG30564
ID ABG30564 standard; Protein; 1286 AA.
XX
AC ABG30564;
XX
DT 07-OCT-2002 (first entry)
XX
DE Alpha-isomaltosylglucosaccharide synthase #2.
XX
KW Alpha-isomaltosylglucosaccharide synthase; sugar production;
KW cyclic tetrasaccharide; cyclo(-6)-alpha-D-glucopyranosyl(1-3);
KW alpha-isomaltosyltransferase; viscosity modifier; sweetener;
KW heat stabiliser; hydrolysis stabiliser; colour stabiliser; foodstuff;
KW cosmetic; drug composition.
XX
OS Unidentified.
XX
PN WO200255708-A1.
XX
PD 18-JUL-2002.
XX
PF 09-JAN-2002; 2002WO-JP00052.
XX
PR 12-JAN-2001; 2001JP-0005441.
XX
PA (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
XX
PI Kubota M, Maruta K, Yamamoto T, Fukuda S;
XX
DR WPI; 2002-520129/55.
DR N-PSDB; ABK88158.
XX
PT New alpha-isomaltosylglucosaccharide synthase of bacterial origin for
PT the production of cyclic tetrasaccharide gum -
XX
PS Disclosure; Page 128-134; 144pp; Japanese.
XX
CC The invention describes novel microbial polypeptides having
CC alpha-isomaltosylglucosaccharide synthase activity. The proteins are
CC useful for producing a sugar (I) having at least three glucose units
CC connected by alpha-1,4 bonds except for alpha-1,6 at the non-reducing
CC end. The invention also describes a method for the production of the
CC cyclic tetrasaccharide cyclo(-6)-alpha-D-glucopyranosyl(1-3)-alpha-D-
CC glucopyranosyl(1-6)-alpha-D-glucopyranosyl(1-3)-alpha-D-glucopyranosyl
CC (1-) (II) by treatment of (I) with alpha-isomaltosyltransferase.
CC (II) and similar sugars in crystalline or syrup form are used as
CC a viscosity modifier, sweetener, heat stabiliser, hydrolysis stabiliser
CC and colour stabiliser for foodstuffs, cosmetics and drug compositions.
CC This is the amino acid sequence of an alpha-isomaltosylglucosaccharide
CC synthase.

SQ	Sequence	1286 AA;	
	Query Match	100.0%;	Score 44; DB 23; Length 1286;
	Best Local Similarity	100.0%;	Pred. No. 3.7;
	Matches	9; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	1 HVSAIIGNLL 9		
DB	36 HVSAIIGNLL 44		
RESULT 4			
ABG30551			
ID	ABG30551 standard; Peptide; 8 AA.		
XX			
AC	ABG30551;		
XX			
DT	07-OCT-2002 (first entry)		
XX			
DE	Alpha-isomaltosylglucosaccharide synthase related peptide #12.		
XX			
KW	Alpha-isomaltosylglucosaccharide synthase; sugar production;		
KW	cyclic tetrasaccharide; cyclo(-6)-alpha-D-glucopyranosyl(1-3);		
KW	alpha-isomaltosyltransferase; viscosity modifier; sweetener;		
KW	heat stabiliser; hydrolysis stabiliser; colour stabiliser; foodstuff;		
KW	cosmetic; drug composition.		
XX			
OS	Unidentified.		
XX			
PN	WO200255708-A1.		
XX			
PD	18-JUL-2002.		
XX			
PF	09-JAN-2002; 2002WO-JP00052.		
XX			
PR	12-JAN-2001; 2001JP-0005441.		
XX			
PA	(HAYB) HAYASHIBARA SEIBUTSU KAGAKU.		
XX			
PI	Kubota M, Maruta K, Yamamoto T, Fukuda S;		
XX			
DR	WPI; 2002-520129/55.		
XX			
PT	New alpha-isomaltosylglucosaccharide synthase of bacterial origin for		
PT	the production of cyclic tetrasaccharide gum -		
XX			
PS	Disclosure; Page 126; 144pp; Japanese.		
XX			
CC	The invention describes novel microbial polypeptides having		
CC	alpha-isomaltosylglucosaccharide synthase activity. The proteins are		
CC	useful for producing a sugar (I) having at least three glucose units		
CC	connected by alpha-1,4 bonds except for alpha-1,6 at the non-reducing		
CC	end. The invention also describes a method for the production of the		
CC	cyclic tetrasaccharide cyclo(-6)-alpha-D-glucopyranosyl(1-3)-alpha-D-		
CC	glucopyranosyl(1-6)-alpha-D-glucopyranosyl(1-3)-alpha-D-glucopyranosyl		
CC	(1-) (II) by treatment of (I) with alpha-isomaltosyltransferase.		
CC	(II) and similar sugars in crystalline or syrup form are used as		
CC	a viscosity modifier, sweetener, heat stabiliser, hydrolysis stabiliser		
CC	and colour stabiliser for foodstuffs, cosmetics and drug compositions.		
CC	This is the amino acid sequence of an alpha-isomaltosylglucosaccharide		
CC	synthase related peptide.		
XX			
SQ	Sequence	8 AA;	
	Query Match	90.9%;	Score 40; DB 23; Length 8;
	Best Local Similarity	100.0%;	Pred. No. 9.3e+05;
	Matches	8; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	1 HVSAIIGNLL 8		
DB	1 HVSAIIGNLL 8		
RESULT 5			
AAG92260			
ID	AAG92260 standard; Protein; 330 AA.		
XX			
AC	AAG92260;		
XX			
DT	26-SEP-2001 (first entry)		
XX			
DE	C glutamicum protein fragment SEQ ID NO: 6014.		
XX			
KW	Coryneform bacterium; amino acid synthesis; vitamin; saccharide;		
KW	organic acid synthesis.		
XX			
OS	Corynebacterium glutamicum.		
XX			
PN	EP1108790-A2.		
XX			
PD	20-JUN-2001.		
XX			
PF	18-DEC-2000; 2000EP-0127688.		
XX			
PR	16-DEC-1999; 99JP-0377484.		
PR	07-APR-2000; 2000JP-0159162.		
PR	03-AUG-2000; 2000JP-0280988.		
XX			
PA	(KYOW) KYOWA HAKKO KOGYO KK.		
XX			
PI	Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;		
PI	Tateishi N, Senoh A, Ikeda M, Ozaki A;		
XX			
DR	WPI; 2001-376931/40.		
DR	N-PSDB; AAH67479.		
XX			
PT	Novel polynucleotides derived from Coryneform bacteria, for identifying		
PT	mutation point of a gene, measuring expression of a gene, analysing		
PT	expression profile or pattern of a gene and identifying homologous gene		
XX			
PS	Claim 17; SEQ ID NO: 6014; 246pp + Sequence Listing; English.		
XX			
CC	The present invention provides a number of nucleotide and protein		
CC	sequences from the Coryneform bacterium Corynebacterium glutamicum. These		
CC	are useful for identifying the mutation point of a gene derived from a		
CC	mutant of coryneform bacterium, measuring expression amount and		
CC	analysing the expression profile or expression pattern of a gene derived		
CC	from Coryneform bacterium, and identifying a homologue of a gene derived		
CC	from Coryneform bacterium. Coryneform bacteria are useful for producing		
CC	amino acids, nucleic acids, vitamins, saccharides and organic acids,		
CC	particularly L-lysine. The present sequence is a protein described		
CC	in the exemplification of the invention.		
CC	Note: The sequence data for this patent did not form part of the printed		
CC	specification, but was obtained in electronic format directly from the		
CC	European Patent Office.		
XX			
SQ	Sequence	330 AA;	
	Query Match	84.1%;	Score 37; DB 22; Length 330;
	Best Local Similarity	87.5%;	Pred. No. 23;
	Matches	7; Conservative	1; Mismatches 0; Indels 0; Gaps 0;
QY	1 HVSAIIGNLL 8		
DB	300 HVSAIIGNLL 307		
RESULT 6			
AAB80034			
ID	AAB80034 standard; Protein; 380 AA.		
XX			
AC	AAB80034;		
XX			
DT	30-APR-2001 (first entry)		
XX			

PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134321.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134370.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0139119.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.

PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.

PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 79.5%; Score 35; DB 21; Length 162;

Best Local Similarity 77.8%; Pred. No. 28;

Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 HVSAIGNLL 9

DB 104 HVELIGNLL 112

RESULT 8

AAG33438

ID AAG33438 standard; Protein; 196 AA.

XX AC AAG33438;

XX DT 18-OCT-2000 (first entry)

XX DE Zea mays protein fragment SEQ ID NO: 40514.

XX KW Protein identification; signal transduction pathway; metabolic pathway;
 XX KW hybridisation assay; genetic mapping; gene expression control; promoter;
 XX KW termination sequence; corn.

XX OS Zea mays subsp. mays.

XX PN EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

XX PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 28-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 28-APR-1999; 99US-0130891.

PR 30-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 04-MAY-1999; 99US-0132407.

PR 05-MAY-1999; 99US-0132484.

PR 06-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR 07-MAY-1999; 99US-0132487.

PR 11-MAY-1999; 99US-0132863.

PR 11-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.
 PR 14-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 02-JUL-1999; 99US-0142154.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 20-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 23-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.

```
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148584.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149829.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152163.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157523.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.

PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 79.5%; Score 35; DB 21; Length 196;
Best Local Similarity 77.8%; Pred. No. 34;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 HVSAIGNLL 9
Db 138 HVLLGNLL 146

RESULT 9
ABU07377
ID ABU07377 standard; Protein; 3931 AA.
XX
AC ABU07377;
XX
DT 28-JAN-2003 (first entry)
XX
DE Human protein NOV9.
XX
KW Human; NOVX; cardiomyopathy; atherosclerosis; cancer;
KW hypertension; diabetes; inflammation; autoimmune disorder; allergy;
KW blood disorder; acquired immunodeficiency syndrome; AIDS; obesity;
KW asthma; immunoglobulin A nephropathy; cirrhosis; arthritis;
KW Alzheimer's disease; Parkinson's disease; goitre; infection; stroke;
KW muscular dystrophy; epilepsy; wasting disorder; neurogenesis;
KW cell differentiation; cell proliferation; haematopoiesis; wound healing;
KW angiogenesis; gene therapy; single nucleotide polymorphism; SNP.
XX
OS Homo sapiens.
XX
XX WO200285922-A2.
XX
XX 31-OCT-2002.
XX
XX 11-APR-2002; 2002WO-US11634.
XX
XX 23-APR-2001; 2001US-285748P.
XX
XX 24-APR-2001; 2001US-286068P.
XX
XX 25-APR-2001; 2001US-286292P.
XX
XX 03-MAY-2001; 2001US-288334P.
XX
XX 16-MAY-2001; 2001US-291241P.
XX
XX 14-SEP-2001; 2001US-322284P.
XX
XX (CURA-) CURAGEN CORP.
XX
XX (MILL-) MILLENNIUM PHARM INC.
XX
XX Pena CEA, Guo X, Shinkets RA, Padigaru M, Kekuda R, Spytek KA;
XX Mehrahan F, Topper JN, Malyankar UM, Wasserman S, Edinger S;
XX Smithson G, Gunther E, Komuves L;
XX
XX WPI: 2003-058712/05.
XX
XX N-PSDB; ABX10231.
XX
XX New NOVX polypeptides and nucleic acids, useful for preventing or
XX treating NOVX-associated disorders, e.g. cancer, cardiomyopathy,
XX atherosclerosis or diabetes, and in chromosome mapping, tissue typing
XX or pharmacogenomics
XX
XX Claim 1; Page 65-66; 301pp; English.
XX
XX The invention relates to an isolated polypeptide comprising any one of
```

CC 17 human NOVX (1-9, 10a, 10b, 11-16) appearing as ABU07369-ABU07385, a
 CC mature form of it, or a variant of them, where one or more residues of
 CC the variant differs in not more than 15 % from the residues of the
 CC sequence of them and their encoding polynucleotides appearing as
 CC ABX10223-ABX10239. Also included are NOVX expression vectors,
 CC transformed cells, antibodies, identifying an agent that binds to or
 CC modulates the expression or activity of NOVX and screening for a
 CC modulator of activity or of latency or predisposition to a NOVX-
 CC associated disorder. The NOVX polypeptides, polynucleotides and
 CC antibodies are useful in manufacturing a medicament for treating or
 CC preventing a syndrome associated with NOVX-associated disorder, such as
 CC cardiomyopathy, atherosclerosis, cancer, hypertension, diabetes,
 CC inflammation, autoimmune disorders, allergies, blood disorders,
 CC acquired immunodeficiency syndrome (AIDS), obesity, asthma,
 CC immunoglobulin (Ig)A nephropathy, cirrhosis, arthritis, Alzheimer's
 CC disease, Parkinson's disease, goitre, infections (e.g. bacterial,
 CC viral, parasitic), stroke, muscular dystrophy, epilepsy, and other
 CC wasting disorders associated with chronic diseases. The nucleic acids
 CC and polypeptides may also be used as targets for the identification of
 CC small molecules that modulate or inhibit e.g. neurogenesis, cell
 CC differentiation, cell proliferation, haematopoiesis, wound healing and
 CC angiogenesis, in gene therapy, in generation of antibodies that bind
 CC immunospecifically to NOVX substances for use in therapeutic or
 CC diagnostic methods. The nucleic acids are further used as hybridisation
 CC probes, in chromosome mapping, tissue typing, preventive medicine, and
 CC pharmacogenomics. The polypeptides are also useful as vaccines.
 CC The present sequence represents a NOVX protein.

SQ Sequence 3931 AA;

Query Match 79.5%; Score 35; DB 24; Length 3931;
 Best Local Similarity 77.8%; Pred. No. 8.8e+02;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 HVSAALGNLL 9
 |||||:
 DB 920 HVSAGVGRLL 928

RESULT 10
 AAG90316
 ID AAG90316 standard; Protein; 109 AA.

XX AC AAG90316;

XX DT 26-SEP-2001 (first entry)

XX DE C glutamicum protein fragment SEQ ID NO: 4070.

XX DE Corynebacterium; amino acid synthesis; vitamin; saccharide;
 KW organic acid synthesis.

XX OS Corynebacterium glutamicum.

XX PN EP1108790-A2.

XX PD 20-JUN-2001.

XX PF 18-DEC-2000; 2000EP-0127698.

XX PR 16-DEC-1999; 99JP-0377484.

XX PR 07-APR-2000; 2000JP-0159162.

XX PR 03-AUG-2000; 2000JP-0280988.

XX PA (KYOW) KYOWA HAKKO KOGYO KK.

XX PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;

XX PI Tateishi N, Senoh A, Ikeda M, Ozaki A;

XX XX WPI; 2001-376931/40.

XX DR N-PSDB; AAH65535.

XX PT Novel polynucleotides derived from Corynebacterium, for identifying

PT mutation point of a gene, measuring expression of a gene, analysing
 PT expression profile or pattern of a gene and identifying homologous gene
 PT -
 XX Claim 17; SEQ ID NO: 4070; 246pp + Sequence Listing; English.
 XX The present invention provides a number of nucleotide and protein
 CC sequences from the Corynebacterium glutamicum. These
 CC are useful for identifying the mutation point of a gene derived from a
 CC mutant of corynebacterium, measuring expression amount and
 CC analysing the expression profile or expression pattern of a gene derived
 CC from Corynebacterium, and identifying a homologue of a gene derived
 CC from corynebacterium. Corynebacterium bacteria are useful for producing
 CC amino acids, nucleic acids, vitamins, saccharides and organic acids,
 CC particularly L-lysine. The present sequence is a protein described
 CC in the exemplification of the invention.
 CC Note: the sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from the
 CC European Patent Office.

XX SQ Sequence 109 AA;

Query Match 77.3%; Score 34; DB 22; Length 109;
 Best Local Similarity 75.0%; Pred. No. 29;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 HVSAALGNL 8
 |||||:
 DB 26 HISAALGSL 33

RESULT 11
 ABJ25590
 ID ABJ25590 standard; Protein; 526 AA.

XX AC ABJ25590;

XX DT 16-APR-2003 (first entry)

XX DE Aspergillus fumigatus essential gene protein #248.

XX KW Fungicide; cytostatic; essential gene; Aspergillus fumigatus; infection;
 KW cancer; contamination; biofilm; antibody; immune response.

XX OS Aspergillus fumigatus.

XX PN WO200286090-A2.

XX PD 31-OCT-2002.

XX PF 23-APR-2002; 2002WO-US13142.

XX PR 23-APR-2001; 2001US-285697P.

XX PR 27-APR-2001; 2001US-287066P.

XX PR 05-JUN-2001; 2001US-295890P.

XX PR 09-JUL-2001; 2001US-303899P.

XX PR 31-AUG-2001; 2001US-316362P.

XX PA (ELIT-) ELITRA PHARM INC.

XX PI Jiang B, Tishkoff D, Zamudio C, Eroshkin AM, Hu W, Lemieux SM;

XX DR WPI; 2003-093124/08.

XX PT New purified or isolated nucleic acids of essential genes of
 PT Aspergillus fumigatus, useful for treating or preventing infections by
 PT A. fumigatus, or for treating a non-infectious disease in a subject
 PT e.g. cancer -

XX PS Disclosure; Page -; 175pp; English.

XX CC The invention relates to novel purified or isolated nucleic acids of
 CC essential genes of Aspergillus fumigatus. The isolated nucleic acids of

CC the invention are used to treat or prevent infections by a pathogenic
 CC organism such as *A. fumigatus*, to treat a non-infectious disease in a
 CC subject (e.g. cancer), to prevent or contain contamination of an object
 CC by *A. fumigatus*, or to prevent or inhibit formation on a surface of a
 CC biofilm comprising *A. fumigatus*. The polynucleotides are useful for
 CC expressing recombinant protein for characterisation, screening or
 CC therapeutic use, as markers for host tissues in which the pathogenic
 CC organisms invade or reside, for comparing with the DNA sequence of *A.*
 CC *fumigatus* to identify duplicated genes or paralogues having the same or
 CC similar biochemical activity and/or function, for comparing with DNA
 CC sequences of other related or distant pathogenic organisms to identify
 CC potential orthologous essential or virulence genes, for selecting and
 CC making oligomers for attachment to a nucleic acid array for examination
 CC of expression patterns, for raising anti-protein antibodies, as an
 CC antigen to raise anti-DNA antibodies or to elicit another immune
 CC response, and for identifying polynucleotides encoding the other protein
 CC with which binding occurs or to identify inhibitors of the binding
 CC interaction. The polypeptides may be used to raise antibodies or to
 CC elicit immune response, as a reagent in assays designed to quantitatively
 CC determine levels of the protein in biological fluids, as a marker for
 CC host tissues in which pathogenic organism invade or reside, and to
 CC isolate correlative receptors or ligands in the case of virulence
 CC factors. This sequence represents a protein of one of the essential genes
 CC of *Aspergillus fumigatus* of the invention.

XX SQ Sequence 526 AA;
 Query Match 77.1%; Score 34; DB 24; Length 526;
 Best Local Similarity 55.8%; Pred. No. 1.6e+02;
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 HVSAAGNLL 9
 : : : : :
 Db 66 HIAAGNLI 74

RESULT 12
 ID ABB84268
 XX ABB84268 standard; peptide; 9 AA.
 AC ABB84268;
 XX
 DT 08-OCT-2002 (first entry)
 XX
 DE B. globisporus dehydration-associated peptide SEQ ID 1.
 KW Dehydration; cyclo-alpha-D-glucopyranosyl; pharmaceutical; agrochemical;
 KW toletry; non-reducing saccharide.
 XX Bacillus globisporus.
 XX WO200257011-A1.
 XX
 PD 25-JUL-2002.
 XX
 PF 17-JAN-2002; 2002WO-JP00288.
 XX
 PR 19-JAN-2001; 2001JP-0010991.
 XX
 PA (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
 XX Kubota M, Nishimoto T, Aga H, Fukuda S, Miyake T;
 PI
 DR WPI; 2002-590706/63.
 XX
 PT Dehydrating agent used for pharmaceuticals, agrochemicals and
 PT toiletries comprises cyclic tetra-saccharide -
 PS Disclosure; Page 134; 140pp; Japanese.
 XX

CC This invention describes a novel dehydrating agent comprising
 CC cyclo(6)-alpha-D-glucopyranosyl-(1-3)-alpha-D-glucopyranosyl-(1-6)-alpha-
 CC D-glucopyranosyl-(1-3)-alpha-glycopyranosyl. The agent of the invention

CC can be used in pharmaceuticals, agrochemicals and toiletries. The agent
 CC is a non-reducing saccharide and dries articles without causing a
 CC deterioration in quality of the articles. This sequence represents a
 CC peptide described in the disclosure of the invention.

XX SQ Sequence 9 AA;
 Query Match 75.0%; Score 33; DB 23; Length 9;
 Best Local Similarity 66.7%; Pred. No. 9.3e+05;
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 HVSAAGNLL 9
 : : : : :
 Db 1 YVSSAGNLI 9

RESULT 13
 ID ABB30540
 XX ABB30540 standard; Peptide; 9 AA.
 AC ABB30540;
 XX
 DT 07-OCT-2002 (first entry)
 XX
 DE Alpha-isomaltosylglucosaccharide synthase related peptide #1.
 KW Alpha-isomaltosylglucosaccharide synthase; sugar production;
 KW cyclic tetrasaccharide; cyclo(-6)-alpha-D-glucopyranosyl(1-3);
 KW alpha-isomaltosyltransferase; viscosity modifier; sweetener;
 KW heat stabiliser; hydrolysis stabiliser; colour stabiliser; foodstuff;
 KW cosmetic; drug composition.

XX Unidentified.
 XX WO200255708-A1.
 XX
 PD 18-JUL-2002.

XX
 PF 09-JAN-2002; 2002WO-JP00052.
 XX
 PR 12-JAN-2001; 2001JP-0005441.

XX (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.

XX Kubota M, Maruta K, Yamamoto T, Fukuda S;

XX WPI; 2002-520129/55.

XX New alpha-isomaltosylglucosaccharide synthase of bacterial origin for
 XX the production of cyclic tetrasaccharide gum -

XX Disclosure; Page 117; 144pp; Japanese.

CC The invention describes novel microbial polypeptides having
 CC alpha-isomaltosylglucosaccharide synthase activity. The proteins are
 CC useful for producing a sugar (I) having at least three glucose units
 CC connected by alpha-1,4 bonds except for alpha-1,6 at the non-reducing
 CC end. The invention also describes a method for the production of the
 CC cyclic tetrasaccharide cyclo(-6)-alpha-D-glucopyranosyl(1-3)-alpha-D-
 CC glucopyranosyl(1-6)-alpha-D-glucopyranosyl(1-3)-alpha-D-glucopyranosyl
 CC (I-) (II) by treatment of (I) with alpha-isomaltosyltransferase,
 CC (II) and similar sugars in crystalline or syrup form are used as
 CC a viscosity modifier, sweetener, heat stabiliser, hydrolysis stabiliser
 CC and colour stabiliser for foodstuffs, cosmetics and drug compositions.
 CC This is the amino acid sequence of an alpha-isomaltosylglucosaccharide
 CC synthase related peptide.

XX SQ Sequence 9 AA;

Query Match 75.0%; Score 33; DB 23; Length 9;
 Best Local Similarity 66.7%; Pred. No. 9.3e+05;
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 HVSAAGNLL 9
 :||:||||:
 Db 1 YVSSAGNLI 9

RESULT 14
 AA017330
 ID AA017330 standard; peptide; 9 AA.
 XX
 AC AA017330;
 XX
 DT 08-JUL-2002 (first entry)
 XX
 DE Alpha-isomaltosylglucosaccharide synthase related peptide #1.
 XX
 KW Alpha-isomaltosylglucosaccharide synthase; enzyme; sweetener; food;
 KW drink; cosmetics; pharmaceutical; cyclic tetrasaccharide.
 XX
 OS Bacillus globisporus.
 XX
 XX W0200210361-A1.
 PN
 XX
 PD 07-FEB-2002.
 XX
 XX 25-JUL-2001; 2001WO-JP06412.
 PF
 XX
 XX 01-AUG-2000; 2000JP-0233364.
 PR
 PR 02-AUG-2000; 2000JP-0234937.
 XX
 XX (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
 PA
 XX
 PI Kubota M, Tsueaki K, Higashiyama T, Fukuda S, Miyake T;
 XX
 XX WPI; 2002-315302/35.
 DR
 XX
 XX Alpha-Isomaltosylglucosaccharide synthase capable of transferring
 PT Alpha-glucosyl from saccharide, useful in producing cyclic
 PT tetrasaccharides and related carbohydrates industrially for application
 PT in compositions e.g. as drugs -
 XX
 PS Claim 5; Page 200; 209pp; Japanese.
 XX
 CC The present invention relates to an alpha-isomaltosylglucosaccharide
 CC synthase which is capable of transferring alpha-glucosyl from a
 CC saccharide to form a specific saccharide which carries an alpha-1,6
 CC glucosyl linkage at its non-reducing end and an alpha-1,4 glucosyl
 CC linkage at ends other than the non-reducing end and has a degree of
 CC glucose polymerization of at most 3, but without substantially elevating
 CC the reducing ability. The synthase is useful in producing cyclic
 CC tetrasaccharides and related carbohydrates industrially for application
 CC in compositions e.g. as sweeteners, low calorie food materials, taste
 CC improvers, flavour improvers, quality improvers, water-separation
 CC inhibitors, stabilizers, excipients, binding agents and pulverisation
 CC bases, especially in foods, drinks, cosmetics and pharmaceuticals. The
 CC present sequence is a peptide described in the exemplification of the
 CC invention.
 XX
 SQ Sequence 9 AA;
 Query Match 75.0%; Score 33; DB 23; Length 9;
 Best Local Similarity 66.7%; Pred. No. 9.3e+05;
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 HVSAAGNLL 9
 :||:||||:
 Db 1 YVSSAGNLI 9

RESULT 15
 AA057279
 ID AA057279 standard; Protein; 188 AA.
 XX
 AC AA057279;

XX 27-FEB-2002 (first entry)
 DT
 XX
 DE Propionibacterium acnes immunogenic protein #18175.
 XX
 KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant.
 XX
 OS Propionibacterium acnes.
 XX
 PN W0200181581-A2.
 XX
 PD 01-NOV-2001.
 XX
 PF 20-APR-2001; 2001WO-US12865.
 XX
 PR 21-APR-2000; 2000US-199047P.
 PR 02-JUN-2000; 2000US-208841P.
 PR 07-JUL-2000; 2000US-216747P.
 XX
 XX (CORI-) CORIXA CORP.
 PA
 XX
 PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
 PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
 XX
 XX WPI; 2001-616774/71.
 DR
 DR N-PSDB; AAS9581.
 XX
 XX Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris -
 PT
 XX
 XX Example 1; SEQ ID No 18474; 1069pp; English.
 CC
 CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
 CC polypeptides. The proteins and their associated DNA sequences are used in
 CC the treatment, prevention and diagnosis of medical conditions caused by
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
 CC P. acnes is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of P. acnes in a patient comprises contacting a
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for P. acnes proteins. These antibodies can be used to
 CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA).
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 188 AA;
 Query Match 75.0%; Score 33; DB 22; Length 188;
 Best Local Similarity 55.6%; Pred. No. 85;
 Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 HVSAAGNLL 9
 :||:||||:
 Db 39 HMTALGNVM 47

Search completed: January 20, 2004, 16:52:46
 Job time : 34.8015 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 20, 2004, 16:53:13 ; Search time 10.5802 Seconds
(without alignments)
35.992 Million cell updates/sec

Title: US-10-089-549-11
Perfect score: 44
Sequence: 1 HVSAAGNLL 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/1aa/5A-COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B-COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A-COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B-COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCTUS-COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	33	75.0	194	3	US-08-959-004-9
2	33	75.0	228	4	US-09-252-991A-26434
3	32	72.7	320	4	US-09-252-991A-23195
4	32	72.7	1976	3	US-09-024-020B-9
5	32	72.7	1976	4	US-09-425-043-9
6	32	72.7	1978	3	US-09-024-020B-3
7	32	72.7	1978	4	US-09-425-043-3
8	32	72.7	1988	3	US-09-024-020B-4
9	32	72.7	1988	4	US-09-425-043-4
10	31	70.5	344	4	US-09-252-991A-28141
11	31	70.5	269	4	US-08-311-731A-38
12	31	70.5	594	4	US-09-252-991A-25988
13	30	68.2	221	4	US-09-252-991A-18014
14	30	68.2	269	4	US-09-252-991A-31516
15	30	68.2	412	2	US-08-463-081B-14
16	30	68.2	412	2	US-08-461-379A-14
17	30	68.2	412	2	US-08-462-390B-14
18	30	68.2	412	3	US-08-463-074B-14
19	30	68.2	412	3	US-08-465-585C-14
20	30	68.2	412	3	US-08-652-446-14
21	30	68.2	412	4	US-09-462-624-2
22	30	68.2	441	4	US-09-328-352-8031
23	30	68.2	585	4	US-09-252-991A-32406
24	30	68.2	659	4	US-09-189-462-4
25	30	68.2	659	4	US-09-196-270-5
26	29	65.9	113	3	US-08-406-030A-29
27	29	65.9	155	3	US-08-959-004-3

28	29	65.9	159	4	US-09-328-352-7798	Sequence 7798, Ap
29	29	65.9	166	4	US-08-679-493A-209	Sequence 209, Appl
30	29	65.9	237	1	US-08-750-532-18	Sequence 18, Appl
31	29	65.9	260	4	US-09-796-774-1	Sequence 1, Appli
32	29	65.9	260	4	US-08-663-831-1	Sequence 1, Appli
33	29	65.9	260	4	US-08-663-831-2	Sequence 2, Appli
34	29	65.9	260	4	US-08-663-831-3	Sequence 3, Appli
35	29	65.9	260	4	US-08-663-831-4	Sequence 4, Appli
36	29	65.9	260	4	US-08-663-831-5	Sequence 5, Appli
37	29	65.9	260	4	US-08-663-831-6	Sequence 6, Appli
38	29	65.9	260	4	US-08-663-831-9	Sequence 9, Appli
39	29	65.9	260	4	US-08-663-831-10	Sequence 10, Appl
40	29	65.9	260	4	US-08-663-831-11	Sequence 11, Appl
41	29	65.9	260	4	US-08-663-831-12	Sequence 12, Appl
42	29	65.9	260	4	US-08-663-831-13	Sequence 13, Appl
43	29	65.9	260	4	US-08-663-831-14	Sequence 14, Appl
44	29	65.9	260	4	US-08-663-831-15	Sequence 15, Appl
45	29	65.9	260	4	US-08-663-831-16	Sequence 16, Appl

ALIGNMENTS

RESULT 1
US-08-959-004-9
; Sequence 9, Application US/08959004
; Patent No. 6197543
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Yue, Henry
; APPLICANT: Corley, Neil C.
; APPLICANT: Lal, Preeti
; APPLICANT: Shah, Purvi
; APPLICANT: Kaser, Matthew
; TITLE OF INVENTION: HUMAN VESICLE MEMBRANE PROTEIN-LIKE
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/959,004
; FILING DATE: Herewith
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0414 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 194 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 297437

US-08-959-004-9

Query Match 75.0%; Score 33; DB 3; Length 194;
Best Local Similarity 87.5%; Pred. No. 30;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 VSALGNLL 9
:|||||
DB 44 LSALGNLL 51

RESULT 2

US-09-252-991A-26434
; Sequence 26434, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 26434
; LENGTH: 228
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26434

Query Match 75.0%; Score 33; DB 4; Length 228;
Best Local Similarity 66.7%; Pred. No. 36;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 HVSALGNLL 9
:|||||
DB 130 HIEALGTL 138

RESULT 3

US-09-252-991A-23195
; Sequence 23195, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 23195
; LENGTH: 320
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23195

Query Match 72.7%; Score 32; DB 4; Length 320;
Best Local Similarity 100.0%; Pred. No. 85;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 SALGNLL 9
:|||||
DB 295 SALGNLL 301

RESULT 4

US-09-024-020B-9
; Sequence 9, Application US/09024020B
; Patent No. 6030810
; GENERAL INFORMATION:
; APPLICANT: DELGADO, STEPHEN G.
; APPLICANT: DIETRICH, PAUL S.
; APPLICANT: FISH, LINDA M.
; APPLICANT: HERMAN, RONALD C.
; APPLICANT: SANGAMESWARAN, LAKSHMI
; TITLE OF INVENTION: NOVEL CLONED TETRODOTOXIN-SENSITIVE
; TITLE OF INVENTION: SODIUM CHANNEL I-SUBUNIT AND A SPLICE VARIANT THEREOF
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: JANET PAULINE CLARK
; STREET: 3401 HILLVIEW AVENUE, MS A2-250
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 94304-1397
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/024,020B
; FILING DATE: 16-FEB-1998
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/039,447
; FILING DATE: 26-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: CLARK, JANET P.
; REGISTRATION NUMBER: 34,799
; REFERENCE/DOCKET NUMBER: R0020B-REG
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 852-3097
; TELEFAX: (650) 855-5322
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1976 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-024-020B-9

Query Match 72.7%; Score 32; DB 3; Length 1976;
Best Local Similarity 66.7%; Pred. No. 6.5e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 HVSALGNLL 9
:|||||
DB 779 HVLAVGNLV 787

RESULT 5

US-09-425-043-9
; Sequence 9, Application US/09425043
; Patent No. 6335172
; GENERAL INFORMATION:
; APPLICANT: DELGADO, STEPHEN G.
; APPLICANT: DIETRICH, PAUL S.
; APPLICANT: FISH, LINDA M.
; APPLICANT: HERMAN, RONALD C.
; APPLICANT: SANGAMESWARAN, LAKSHMI
; TITLE OF INVENTION: NOVEL CLONED TETRODOTOXIN-SENSITIVE
; TITLE OF INVENTION: SODIUM CHANNEL I-SUBUNIT AND A SPLICE VARIANT THEREOF
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: JANET PAULINE CLARK
; STREET: 3401 HILLVIEW AVENUE, MS A2-250

;/ CITY: PALO ALTO
;/ STATE: CA
;/ COUNTRY: U.S.A.
;/ ZIP: 94304-1397
;/ COMPUTER READABLE FORM:
;/ MEDIUM TYPE: Floppy disk
;/ OPERATING SYSTEM: IBM PC compatible
;/ SOFTWARE: Patent In Release #1.0, Version #1.30
;/ CURRENT APPLICATION DATA:
;/ APPLICATION NUMBER: US/09/425,043
;/ FILING DATE:
;/ CLASSIFICATION:
;/ PRIOR APPLICATION DATA:
;/ APPLICATION NUMBER: US 09/024,020
;/ FILING DATE: 16-FEB-1998
;/ APPLICATION NUMBER: US 60/039,447
;/ FILING DATE: 26-FEB-1997
;/ ATTORNEY/AGENT INFORMATION:
;/ NAME: CLARK, JANET P.
;/ REGISTRATION NUMBER: 34,799
;/ REFERENCE/DOCKET NUMBER: R0020B-REG
;/ TELECOMMUNICATION INFORMATION:
;/ TELEPHONE: (650) 852-3097
;/ TELEFAX: (650) 855-5322
;/ INFORMATION FOR SEQ ID NO: 9:
;/ SEQUENCE CHARACTERISTICS:
;/ LENGTH: 1976 amino acids
;/ TYPE: amino acid
;/ STRANDEDNESS: single
;/ TOPOLOGY: linear
;/ MOLECULE TYPE: peptide
;/ US-09-425-043-9

Query Match 72.7%; Score 32; DB 4; Length 1976;
Best Local Similarity 66.7%; Pred. No. 6.5e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 HVSAIGNLL 9
Db 779 HVLAUNLV 787

RESULT 6
US-09-024-020B-3
; Sequence 3, Application US/09024020B
; Patent No. 6030810
; GENERAL INFORMATION:
; APPLICANT: DELGADO, STEPHEN G.
; APPLICANT: DIETRICH, PAUL S.
; APPLICANT: FISH, LINDA M.
; APPLICANT: HERMAN, RONALD C.
; APPLICANT: SANGAMESWARAN, LAKSHMI
; TITLE OF INVENTION: NOVEL CLONED TETRODOTOXIN-SENSITIVE
; TITLE OF INVENTION: SODIUM CHANNEL 1-SUBUNIT AND A SPLICE VARIANT THEREOF
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: JANET PAULINE CLARK
; STREET: 3401 HILLVIEW AVENUE, MS A2-250
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 94304-1397
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/024,020B
; FILING DATE: 16-FEB-1998
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:

;/ APPLICATION NUMBER: US 60/039,447
;/ FILING DATE: 26-FEB-1997
;/ ATTORNEY/AGENT INFORMATION:
;/ NAME: CLARK, JANET P.
;/ REGISTRATION NUMBER: 34,799
;/ REFERENCE/DOCKET NUMBER: R0020B-REG
;/ TELECOMMUNICATION INFORMATION:
;/ TELEPHONE: (650) 852-3097
;/ TELEFAX: (650) 855-5322
;/ INFORMATION FOR SEQ ID NO: 3:
;/ SEQUENCE CHARACTERISTICS:
;/ LENGTH: 1978 amino acids
;/ TYPE: amino acid
;/ STRANDEDNESS: single
;/ TOPOLOGY: linear
;/ MOLECULE TYPE: peptide
;/ US-09-024-020B-3

Query Match 72.7%; Score 32; DB 3; Length 1978;
Best Local Similarity 66.7%; Pred. No. 6.5e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 HVSAIGNLL 9
Db 781 HVLAUNLV 789

RESULT 7
US-09-425-043-3
; Sequence 3, Application US/09425043
; Patent No. 6335172
; GENERAL INFORMATION:
; APPLICANT: DELGADO, STEPHEN G.
; APPLICANT: DIETRICH, PAUL S.
; APPLICANT: FISH, LINDA M.
; APPLICANT: HERMAN, RONALD C.
; APPLICANT: SANGAMESWARAN, LAKSHMI
; TITLE OF INVENTION: NOVEL CLONED TETRODOTOXIN-SENSITIVE
; TITLE OF INVENTION: SODIUM CHANNEL 1-SUBUNIT AND A SPLICE VARIANT THEREOF
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: JANET PAULINE CLARK
; STREET: 3401 HILLVIEW AVENUE, MS A2-250
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 94304-1397
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/425,043
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/024,020
; FILING DATE: 16-FEB-1998
; APPLICATION NUMBER: US 60/039,447
; FILING DATE: 26-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: CLARK, JANET P.
; REGISTRATION NUMBER: 34,799
; REFERENCE/DOCKET NUMBER: R0020B-REG
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 852-3097
; TELEFAX: (650) 855-5322
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1978 amino acids
; TYPE: amino acid
; STRANDEDNESS: single

; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-425-043-3

Query Match 72.7%; Score 32; DB 4; Length 1978;
Best Local Similarity 66.7%; Pred. No. 6.5e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 HVSAAGNLL 9
|||:|:|:
Db 781 HVLAVGNLV 789

RESULT 8
US-09-024-020B-4
; Sequence 4, Application US/09024020B
; Patent No. 6030810
; GENERAL INFORMATION:
; APPLICANT: DELGADO, STEPHEN G.
; APPLICANT: DIETRICH, PAUL S.
; APPLICANT: FISH, LINDA M.
; APPLICANT: HERMAN, RONALD C.
; APPLICANT: SANGAMESWARAN, LAKSHMI
; TITLE OF INVENTION: NOVEL CLONED TETRODOTOXIN-SENSITIVE
; TITLE OF INVENTION: SODIUM CHANNEL I-SUBUNIT AND A SPLICE VARIANT THEREOF
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: JANET PAULINE CLARK
; STREET: 3401 HILLVIEW AVENUE, MS A2-250
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 94304-1397
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/024,020B
; FILING DATE: 16-FEB-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/039,447
; FILING DATE: 26-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: CLARK, JANET P.
; REGISTRATION NUMBER: 34,799
; REFERENCE/DOCKET NUMBER: R0020B-REG
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-5322
; TELEFAX: (650) 852-3097
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1988 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-024-020B-4

Query Match 72.7%; Score 32; DB 3; Length 1988;
Best Local Similarity 66.7%; Pred. No. 6.5e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 HVSAAGNLL 9
|||:|:|:
Db 791 HVLAVGNLV 799

RESULT 9
US-09-425-043-4
; Sequence 4, Application US/09425043

; Patent No. 6335172
; GENERAL INFORMATION:
; APPLICANT: DELGADO, STEPHEN G.
; APPLICANT: DIETRICH, PAUL S.
; APPLICANT: FISH, LINDA M.
; APPLICANT: HERMAN, RONALD C.
; APPLICANT: SANGAMESWARAN, LAKSHMI
; TITLE OF INVENTION: NOVEL CLONED TETRODOTOXIN-SENSITIVE
; TITLE OF INVENTION: SODIUM CHANNEL I-SUBUNIT AND A SPLICE VARIANT THEREOF
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: JANET PAULINE CLARK
; STREET: 3401 HILLVIEW AVENUE, MS A2-250
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 94304-1397
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/425,043
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/024,020
; FILING DATE: 16-FEB-1997
; APPLICATION NUMBER: US 60/039,447
; FILING DATE: 26-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: CLARK, JANET P.
; REGISTRATION NUMBER: 34,799
; REFERENCE/DOCKET NUMBER: R0020B-REG
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 852-3097
; TELEFAX: (650) 855-5322
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1988 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-425-043-4

Query Match 72.7%; Score 32; DB 4; Length 1988;
Best Local Similarity 66.7%; Pred. No. 6.5e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 HVSAAGNLL 9
|||:|:|:
Db 791 HVLAVGNLV 799

RESULT 10
US-09-252-991A-28141
; Sequence 28141, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: ABRUINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 28141

```
; LENGTH: 344
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28141

Query Match
Best Local Similarity 70.5%; Score 31; DB 4; Length 344;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 HVSAAGNL 8
   | | | | |
Db 67 HYDALGNL 74

RESULT 11
US-08-311-731A-38
; Sequence 38, Application US/08311731A
; Patent No. 6583266
; GENERAL INFORMATION:
; APPLICANT: SMITH, DOUGLAS
; APPLICANT: MAO, JEN-I
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; TITLE OF INVENTION: RELATING TO MYCOBACTERIUM TUBERCULOSIS AND LAPRAE FOR
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 411
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: WOLF, GREENFIELD & SACKS, P.C.
; STREET: 600 ATLANTIC AVENUE
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02210
; COMPUTER READABLE FORM:
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/311,731A
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: GATES, EDWARD R.
; REGISTRATION NUMBER: 31,616
; REFERENCE/DOCKET NUMBER: C0044/7125
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/720-3500
; TELEFAX: 617/720-2441
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 369 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: MYCOBACTERIUM LEPRAE
US-08-311-731A-38

Query Match
Best Local Similarity 70.5%; Score 31; DB 4; Length 369;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 VSAAGNL 9
   | | | | |
Db 132 IQAAGNL 139

RESULT 12
US-09-252-991A-25988
; Sequence 25988, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25988
; LENGTH: 594
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25988

Query Match
Best Local Similarity 70.5%; Score 31; DB 4; Length 594;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 HVSAAGNL 9
   | | | | |
Db 187 HVEDLGNL 195

RESULT 13
US-09-252-991A-18014
; Sequence 18014, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 18014
; LENGTH: 221
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18014

Query Match
Best Local Similarity 68.2%; Score 30; DB 4; Length 221;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 HVSAAGNL 9
   | | | | |
Db 74 HQRAVGNL 82

RESULT 14
US-09-252-991A-31516
; Sequence 31516, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
```

```
; SEQ ID NO 31516
; LENGTH: 269
; TYPE: PRI
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31516

Query Match      68.2%; Score 30; DB 4; Length 269;
Best Local Similarity 71.4%; Pred. No. 1.8e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy      1 HVSALGN 7
Db      243 HIRALGN 249
      |: |||||

RESULT 15
US-08-463-081B-14
; Sequence 14, Application US/08463081B
; Patent No. 5871960
; Patent No. 5871960 5837487
; GENERAL INFORMATION:
; APPLICANT: Smith, Kendall A. & Beadling, Carol
; TITLE OF INVENTION: Nucleic Acids Encoding CRS Polypeptide,
; TITLE OF INVENTION: Vector and Transformed Cell Thereof, and Expression Thereof
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PRETTY, SCHROEDER & POPLAWSKI
; STREET: 444 South Flower St. - Suite 1900
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0,
; SOFTWARE: Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/463,081B
; FILING DATE: 5-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/104,736
; FILING DATE: 10-AUG-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/796,066
; FILING DATE: 20-NOV-91
; ATTORNEY/AGENT INFORMATION:
; NAME: Viviana Amzel, Ph. D.
; REGISTRATION NUMBER: 30,930
; REFERENCE/DOCKET NUMBER: P66 38150 (DART-060)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 622-7700
; TELEFAX: (213) 489-4210
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 412 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-463-081B-14

Query Match      68.2%; Score 30; DB 2; Length 412;
Best Local Similarity 66.7%; Pred. No. 2.9e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy      1 HVSALGNLL 9
Db      103 HVRALTNLI 111
      |||||
```

Search completed: January 20, 2004, 17:47:11
Job time : 11.5802 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 20, 2004, 16:54:27 ; Search time 23.8397 Seconds
(without alignments)
77.196 Million cell updates/sec

Title: US-10-089-549-11
Perfect score: 44
Sequence: 1 HVSGALGNLL 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 762491 seqs, 204481190 residues

Total number of hits satisfying chosen parameters: 762491

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US05_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	44	100.0	9	12	US-10-089-549-11
2	37	84.1	330	10	US-09-738-626-6014
3	35	79.5	3931	12	US-10-120-801-18
4	34	77.3	109	10	US-09-738-626-4070
5	34	77.3	526	15	US-10-128-714-3248
6	33	75.0	9	12	US-10-089-549-11
7	32	72.7	312	14	US-10-004-717-64
8	32	72.7	407	12	US-10-369-493-3506
9	32	72.7	416	12	US-10-369-493-12762
10	31	70.5	104	9	US-09-864-761-35475
11	31	70.5	174	11	US-09-769-787-113
12	31	70.5	222	10	US-09-738-626-4456
13	31	70.5	225	12	US-10-017-161-728
14	31	70.5	262	15	US-10-156-761-11343
15	31	70.5	355	12	US-10-029-386-32133

16	31	70.5	371	12	US-10-369-493-9320
17	31	70.5	402	12	US-10-368-995-10
18	31	70.5	402	12	US-10-368-995-12
19	31	70.5	428	12	US-10-369-493-9614
20	31	70.5	458	12	US-10-369-493-2945
21	31	70.5	481	12	US-10-369-493-16277
22	31	70.5	494	12	US-10-369-493-15535
23	31	70.5	494	12	US-10-369-493-15906
24	31	70.5	571	12	US-10-369-493-3674
25	31	70.5	685	12	US-09-949-029-68
26	31	70.5	1842	12	US-10-369-493-2225
27	30	68.2	210	15	US-10-156-761-10382
28	30	68.2	265	12	US-10-238-075-467
29	30	68.2	327	12	US-10-369-493-16982
30	30	68.2	331	15	US-10-156-761-13233
31	30	68.2	355	15	US-10-156-761-13820
32	30	68.2	397	12	US-10-369-493-261
33	30	68.2	404	12	US-10-231-913-76
34	30	68.2	411	12	US-10-231-913-77
35	30	68.2	412	10	US-09-950-772-6
36	30	68.2	412	14	US-10-078-650-17
37	30	68.2	412	15	US-10-094-649-2
38	30	68.2	412	15	US-10-289-360-4
39	30	68.2	416	12	US-10-029-386-33130
40	30	68.2	471	12	US-10-246-330-16
41	30	68.2	510	12	US-10-369-493-5737
42	30	68.2	556	15	US-10-128-714-3458
43	30	68.2	556	15	US-10-128-714-3584
44	30	68.2	574	12	US-10-369-493-21968
45	30	68.2	576	15	US-10-128-714-8458

ALIGNMENTS

RESULT 1

US-10-089-549-11
; Sequence 11, Application US/10089549
; Publication No. US20030194762A1
; GENERAL INFORMATION:
; APPLICANT: KUBOTA, Michio
; APPLICANT: TSUSAKI, Keiji
; APPLICANT: HIGASHIYAMA, Takanobu
; APPLICANT: FUKUDA, Shigeharu
; APPLICANT: MIYAKE, Toshio
; TITLE OF INVENTION: ALPHA-ISOMALTOSYLGLUCOSACCHARIDE-FORMING ENZYME, PROCESS AND USES
; FILE REFERENCE: KUBOTA-9
; CURRENT APPLICATION NUMBER: US/10/089,549
; CURRENT FILING DATE: 2002-12-13
; PRIOR APPLICATION NUMBER: JP 233364/2000
; PRIOR FILING DATE: 2000-08-01
; PRIOR APPLICATION NUMBER: JP 234937/2000
; PRIOR FILING DATE: 2000-08-02
; PRIOR APPLICATION NUMBER: PCT/JP01/06412
; PRIOR FILING DATE: 2001-07-25
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Bacillus globisporus
US-10-089-549-11

Query Match 100.0%; Score 44; DB 12; Length 9;
Best Local Similarity 100.0%; Pred. No. 6.7e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HVSGALGNLL 9
DB 1 HVSGALGNLL 9

RESULT 2

US-09-738-626-6014
 ; Sequence 6014, Application US/09738626
 ; Publication No. US20020197605A1
 ; GENERAL INFORMATION:
 ; APPLICANT: NAKAGAWA, SATOSHI
 ; APPLICANT: MIZOGUCHI, HIROSHI
 ; APPLICANT: ANDO, SEIKO
 ; APPLICANT: HAYASHI, MIKIRO
 ; APPLICANT: OCHIAI, KEIKO
 ; APPLICANT: YOKOI, HARUHIKO
 ; APPLICANT: TATEISHI, NAKO
 ; APPLICANT: SENOH, AKIHIRO
 ; APPLICANT: IKEDA, MASATO
 ; APPLICANT: OZAKI, AKIO
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 ; FILE REFERENCE: 249-125
 ; CURRENT APPLICATION NUMBER: US/09/738,626
 ; PRIOR FILING DATE: 2000-12-18
 ; PRIOR APPLICATION NUMBER: JP 99/377484
 ; PRIOR FILING DATE: 1999-12-16
 ; PRIOR APPLICATION NUMBER: JP 00/159162
 ; PRIOR FILING DATE: 2000-04-07
 ; PRIOR APPLICATION NUMBER: JP 00/280988
 ; PRIOR FILING DATE: 2000-08-03
 ; NUMBER OF SEQ ID NOS: 7059
 ; SOFTWARE: PatentIn ver. 3.0
 ; SEQ ID NO 6014
 ; LENGTH: 330
 ; TYPE: PRT
 ; ORGANISM: Corynebacterium glutamicum
 US-09-738-626-6014

Query Match 84.1%; Score 37; DB 10; Length 330;
 Best Local Similarity 87.5%; Pred. No. 25;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 HVSAALGNL 8
 DB 300 HVNALGNL 307

RESULT 3

US-10-120-801-18
 ; Sequence 18, Application US/10120801
 ; Publication No. US20030203843A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Pena, Carol
 ; APPLICANT: Guo, Xiaojia
 ; APPLICANT: Shinkets, Richard
 ; APPLICANT: Padigaru, Muralidhara
 ; APPLICANT: Kekuda, Ramesh
 ; APPLICANT: Spytek, Kimberly
 ; APPLICANT: Mehraban, Fuad
 ; APPLICANT: Topper, James N.
 ; APPLICANT: Malyankar, Uriel
 ; APPLICANT: Wasserman, Scott
 ; APPLICANT: Edinger, Shlomit
 ; APPLICANT: Smithson, Glennda
 ; APPLICANT: Gunther, Erik
 ; APPLICANT: Komuves, Laszlo
 ; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
 ; FILE REFERENCE: 21402-340
 ; CURRENT APPLICATION NUMBER: US/10/120,801
 ; CURRENT FILING DATE: 2002-04-11
 ; PRIOR APPLICATION NUMBER: 60/285748
 ; PRIOR FILING DATE: 2001-04-23
 ; PRIOR APPLICATION NUMBER: 60/286068
 ; PRIOR FILING DATE: 2001-04-24
 ; PRIOR APPLICATION NUMBER: 60/286292
 ; PRIOR FILING DATE: 2001-04-25
 ; PRIOR APPLICATION NUMBER: 60/288334
 ; PRIOR FILING DATE: 2001-05-03

; PRIOR APPLICATION NUMBER: 60/291241
 ; PRIOR FILING DATE: 2001-05-16
 ; PRIOR APPLICATION NUMBER: 60/322284
 ; PRIOR FILING DATE: 2001-09-14
 ; PRIOR APPLICATION NUMBER: 60/285609
 ; PRIOR FILING DATE: 2001-04-20
 ; NUMBER OF SEQ ID NOS: 155
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 18
 ; LENGTH: 3931
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-120-801-18

Query Match 79.5%; Score 35; DB 12; Length 3931;
 Best Local Similarity 77.8%; Pred. No. 8.7e+02;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 HVSAALGNL 9
 DB 920 HVSAVGRLL 928

RESULT 4

US-09-738-626-4070
 ; Sequence 4070, Application US/09738626
 ; Publication No. US20020197605A1
 ; GENERAL INFORMATION:
 ; APPLICANT: NAKAGAWA, SATOSHI
 ; APPLICANT: MIZOGUCHI, HIROSHI
 ; APPLICANT: ANDO, SEIKO
 ; APPLICANT: HAYASHI, MIKIRO
 ; APPLICANT: OCHIAI, KEIKO
 ; APPLICANT: YOKOI, HARUHIKO
 ; APPLICANT: TATEISHI, NAKO
 ; APPLICANT: SENOH, AKIHIRO
 ; APPLICANT: IKEDA, MASATO
 ; APPLICANT: OZAKI, AKIO
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 ; FILE REFERENCE: 249-125
 ; CURRENT APPLICATION NUMBER: US/09/738,626
 ; CURRENT FILING DATE: 2000-12-18
 ; PRIOR APPLICATION NUMBER: JP 99/377484
 ; PRIOR FILING DATE: 1999-12-16
 ; PRIOR APPLICATION NUMBER: JP 00/159162
 ; PRIOR FILING DATE: 2000-04-07
 ; PRIOR APPLICATION NUMBER: JP 00/280988
 ; PRIOR FILING DATE: 2000-08-03
 ; NUMBER OF SEQ ID NOS: 7059
 ; SOFTWARE: PatentIn ver. 3.0
 ; SEQ ID NO 4070
 ; LENGTH: 109
 ; TYPE: PRT
 ; ORGANISM: Corynebacterium glutamicum
 US-09-738-626-4070

Query Match 77.3%; Score 34; DB 10; Length 109;
 Best Local Similarity 75.0%; Pred. No. 30;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 HVSAALGNL 8
 DB 26 HISAALGNL 33

RESULT 5

US-10-128-714-3248
 ; Sequence 3248, Application US/10128714
 ; Publication No. US20030119013A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Jiang, Bo
 ; APPLICANT: Hu, Wenqi
 ; APPLICANT: Tishkoff, Daniel

APPLICANT: Zamudio, Carlos
APPLICANT: Eroshkin, Alexey M
APPLICANT: Lemieux, Sebastien M
TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
TITLE OF INVENTION: Methods of Use
FILE REFERENCE: 10182-018-999
CURRENT APPLICATION NUMBER: US/10/128,714
CURRENT FILING DATE: 2002-04-23
PRIOR APPLICATION NUMBER: US 60/285,697
PRIOR FILING DATE: 2001-04-23
PRIOR APPLICATION NUMBER: US 60/287,066
PRIOR FILING DATE: 2001-04-27
PRIOR APPLICATION NUMBER: US 60/295,890
PRIOR FILING DATE: 2001-06-05
PRIOR APPLICATION NUMBER: US 60/303,899
PRIOR FILING DATE: 2001-07-09
PRIOR APPLICATION NUMBER: US 60/316,362
PRIOR FILING DATE: 2001-08-31
NUMBER OF SEQ ID NOS: 8503
SOFTWARE: PatentIn version 3.1
SEQ ID NO 3248
LENGTH: 526
TYPE: PRT
ORGANISM: Aspergillus fumigatus
US-10-128-714-3248

Query Match 77.3%; Score 34; DB 15; Length 526;

Best Local Similarity 55.6%; Pred. No. 1.6e+02;

Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 HVSAAGNLL 9

Db 66 HIRAGNLI 74

RESULT 6
US-10-089-549-1
Sequence 1, Application US/10089549
Publication No. US20030194762A1
GENERAL INFORMATION:
APPLICANT: KUBOTA, Michio
APPLICANT: TSUSAKI, Keiji
APPLICANT: HIGASHIYAMA, Takanobu
APPLICANT: FUKUDA, Shigeharu
APPLICANT: MIYAKE, Toshio
TITLE OF INVENTION: ALPHA-1SOMALTOXYLGLUCOSACCHARIDE-FORMING ENZYME, PROCESS AND USES
TITLE OF INVENTION: SAME
FILE REFERENCE: KUBOTA=9
CURRENT APPLICATION NUMBER: US/10/089,549
CURRENT FILING DATE: 2002-12-13
PRIOR APPLICATION NUMBER: JP 233364/2000
PRIOR FILING DATE: 2000-08-01
PRIOR APPLICATION NUMBER: JP 234937/2000
PRIOR FILING DATE: 2000-08-02
PRIOR APPLICATION NUMBER: PCT/JP01/06412
PRIOR FILING DATE: 2001-07-25
NUMBER OF SEQ ID NOS: 19
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 9
TYPE: PRT
ORGANISM: Bacillus globisporus
US-10-089-549-1

Query Match 75.0%; Score 33; DB 12; Length 9;

Best Local Similarity 66.7%; Pred. No. 6.7e+05;

Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 HVSAAGNLL 9

Db 1 YVSSAGNLI 9

RESULT 7
US-10-004-717-64
Sequence 64, Application US/10004717
Publication No. US20020192665A1
GENERAL INFORMATION:
APPLICANT: ZOGHEI, HUDA Y.
APPLICANT: YANG, QI
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPEUTIC USE OF AN
TITLE OF INVENTION: ATONAL ASSOCIATED SEQUENCE FOR DEAFNESS,
TITLE OF INVENTION: OSTEOARTHRITIS, AND ABNORMAL CELL PROLIFERATION
FILE REFERENCE: P01899US4
CURRENT APPLICATION NUMBER: US/10/004,717
CURRENT FILING DATE: 2002-08-16
PRIOR APPLICATION NUMBER: 09/585,645
PRIOR FILING DATE: 2000-06-01
PRIOR APPLICATION NUMBER: 60/176,993
PRIOR FILING DATE: 2000-01-19
PRIOR APPLICATION NUMBER: 60/137,060
PRIOR FILING DATE: 1999-06-01
NUMBER OF SEQ ID NOS: 69
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 64
LENGTH: 312
TYPE: PRT
ORGANISM: Drosophila
US-10-004-717-64

Query Match 72.7%; Score 32; DB 14; Length 312;

Best Local Similarity 66.7%; Pred. No. 2.3e+02;

Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 HVSAAGNLL 9

Db 303 YISAGDLL 311

RESULT 8
US-10-369-493-3506
Sequence 3506, Application US/10369493
Publication No. US20030233675A1
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xianfeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369,493
CURRENT FILING DATE: 2003-02-28
PRIOR APPLICATION NUMBER: US 60/360,039
PRIOR FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 3506
LENGTH: 407
TYPE: PRT
ORGANISM: Neurospora crassa
FEATURE:
NAME/KEY: unsure
LOCATION: (1)..(407)
OTHER INFORMATION: unsure at all xaa locations
US-10-369-493-3506

Query Match 72.7%; Score 32; DB 12; Length 407;

Best Local Similarity 75.0%; Pred. No. 3.1e+02;

Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 VSAGNLL 9

Db 127 ISSAGNLL 134

```
RESULT 9
US-10-369-493-12762
; Sequence 12762, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 12762
; LENGTH: 416
; TYPE: PRT
; ORGANISM: Aspergillus nidulans
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)...(416)
; OTHER INFORMATION: unsure at all xaa locations
US-10-369-493-12762

Query Match 72.7%; Score 32; DB 12; Length 416;
Best Local Similarity 87.5%; Pred. No. 3.2e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 VSALGNLL 9
Db 152 VSAPGNLL 159

RESULT 10
US-09-864-761-35475
; Sequence 35475, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
```

```
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 35475
; LENGTH: 104
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL079307.2
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 3
; OTHER INFORMATION: EXPRESSED IN HEAT, SIGNAL = 0.92
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.94
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 33
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.6
; OTHER INFORMATION: EST HUMAN HIT: AA042813.1, EVALUE 2.00e-48
; OTHER INFORMATION: SWISSPROT HIT: Q9S157, EVALUE 2.00e-28
US-09-864-761-35475

Query Match 70.5%; Score 31; DB 9; Length 104;
Best Local Similarity 87.5%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 VSALGNLL 9
Db 21 VMALGNLL 28

RESULT 11
US-09-769-787-113
; Sequence 113, Application US/09769787
; Publication No. US20030091577A1
; GENERAL INFORMATION:
; APPLICANT: Microbial Technics Limited
; APPLICANT: Gilbert, Christophe FG
; APPLICANT: Hansbro, Philip M
; TITLE OF INVENTION: Proteins
; FILE REFERENCE: PWC/P21129WO
; CURRENT APPLICATION NUMBER: US/09/769,787
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: GB 9816337.1
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: US 60/125164
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 388
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 113
; LENGTH: 174
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-769-787-113

Query Match 70.5%; Score 31; DB 11; Length 174;
Best Local Similarity 66.7%; Pred. No. 2e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
```

QY 1 HVSAALGNLL 9
| | | | | :
Db 96 HVSAALGNL 104

RESULT 12

US-09-738-626-4456
; Sequence 4456, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09738, 626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 4456
; LENGTH: 222
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-4456

Query Match 70.5%; Score 31; DB 10; Length 222;
Best Local Similarity 75.0%; Pred. No. 2.6e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 HVSAALGNL 8
| | | | | :
Db 23 HVSAALNV 30

RESULT 13

US-10-017-161-728
; Sequence 728, Application US/10017161
; Publication No. US20030143668A1
; GENERAL INFORMATION:
; APPLICANT: SUMA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABURATANI, HIROYUKI
; TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 084335/0152
; CURRENT APPLICATION NUMBER: US/10/017, 161
; CURRENT FILING DATE: 2002-12-18
; PRIOR APPLICATION NUMBER: JP 2001/246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2430
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 728
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-017-161-728

Query Match 70.5%; Score 31; DB 12; Length 225;
Best Local Similarity 66.7%; Pred. No. 2.6e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 HVSAALGNLL 9
| | | | | :
Db 42 HVSAALGSLI 50

RESULT 14

US-10-156-761-11343
; Sequence 11343, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156, 761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 11343
; LENGTH: 262
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-11343

Query Match 70.5%; Score 31; DB 15; Length 262;
Best Local Similarity 75.0%; Pred. No. 3.1e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 HVSAALGNL 8
| | | | | :
Db 236 HLIALGNL 243

RESULT 15

US-10-029-386-32133
; Sequence 32133, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029, 386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 32133
; LENGTH: 355
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC020667.3
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4.5
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.2
; OTHER INFORMATION: SWISSPROT HIT: P42223, EVALU6 6.00e-01
US-10-029-386-32133

Query Match 70.5%; Score 31; DB 12; Length 355;
Best Local Similarity 62.5%; Pred. No. 4.2e+02;

Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HVSALGNI 8

Db 165 HLNSLGNL 172

Search completed: January 20, 2004, 17:53:14
Job time : 24.8397 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 20, 2004, 16:43:26 ; Search time 9.34351 Seconds
(without alignments)
92.633 Million cell updates/sec

Title: US-10-089-549-11
Perfect score: 44
Sequence: 1 HVSAALGNLL 9
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues
Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 76:.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	34	77.3	203	1 S65033	NADH2 dehydrogenas
2	34	77.3	221	1 S02156	NADH2 dehydrogenas
3	34	77.3	1133	2 T01920	probable RNA-direc
4	33	75.0	149	2 D83079	hypothetical prote
5	33	75.0	194	2 S28850	membrane protein P
6	33	75.0	308	2 T30751	hypothetical prote
7	33	75.0	506	2 S76196	hypothetical prote
8	33	75.0	801	1 A44165	1-phosphatidylinos
9	32	72.7	128	2 S69745	hypothetical prote
10	32	72.7	147	2 D75073	hit-like protein P
11	32	72.7	187	2 T17736	probable superoxid
12	32	72.7	220	2 T49580	probable oligomyci
13	32	72.7	312	2 T35940	probable transport
14	32	72.7	312	2 A40708	basic-helix-loop-h
15	32	72.7	326	2 AE0142	molybdenum cofacto
16	32	72.7	554	2 A70904	probable acid-CoA
17	32	72.7	1180	1 NCECX5	exodeoxyribonuclea
18	32	72.7	1180	2 E91088	DNA helicase RecB
19	32	72.7	1180	2 G85933	DNA helicase RecB
20	32	72.7	1181	2 AB0865	exonuclease V chai
21	32	72.7	1976	2 I56555	sodium channel pro
22	31	70.5	98	2 T42264	hypothetical prote
23	31	70.5	143	1 QOEC32	hypothetical prote
24	31	70.5	174	2 G95083	conserved domain p
25	31	70.5	174	2 C97951	hypothetical prote
26	31	70.5	200	2 F71866	hypothetical prote
27	31	70.5	217	2 D71866	hypothetical prote
28	31	70.5	303	2 E83131	probable transcrip
29	31	70.5	330	1 A87198	phosphoesterase-re

30	31	70.5	356	2 G81907	probable integral
31	31	70.5	356	2 B81107	transporter, proba
32	31	70.5	402	1 A34761	plasma-membrane
33	31	70.5	410	2 C64831	yeaQ protein - Esc
34	31	70.5	410	2 G90753	hypothetical prote
35	31	70.5	410	2 E85617	hypothetical prote
36	31	70.5	442	2 T18507	hypothetical prote
37	31	70.5	458	2 G72377	aspartate ammonia
38	31	70.5	576	2 AB1165	pyruvate oxidase h
39	31	70.5	576	2 AB1524	pyruvate oxidase h
40	31	70.5	586	2 G71656	single-stranded-dn
41	31	70.5	889	2 C86257	resistance to Pseu
42	31	70.5	889	2 T32999	hypothetical prote
43	31	70.5	1045	2 T13481	hypothetical prote
44	31	70.5	1056	2 H84777	probable kinesin-r
45	31	70.5	1842	2 T43409	probable fatty-aci

ALIGNMENTS

RESULT 1

S65033
NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 6 - dermatophytic fungus (Trichophyton rubrum)
C:Species: mitochondrion Trichophyton rubrum
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 03-Jun-2002
C:Accession: S65033
R:de Bievre, C.; Dujon, B.
Curr. Genet. 28, 553-559, 1995
A:Title: Organisation of the mitochondrial genome of Trichophyton rubrum. DNA sequence of the
III gene, the ATPase subunit-8 gene and six tRNA genes that correspond respectively to
A:Reference number: S65031; MUID:96132111; PMID:8593686
A:Accession: S65033
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-203 <DEB>
A:Cross-References: EMBL:X88896; NID:gi1017446; PIDN:CAA61356.1; PID:gi1017449
A>Note: The authors translated the codon CCT for residue 39 as Ala and CCT for residue 1
C:Genetics:
A:Genome: mitochondrion
A:Genetic code: SGC3
C:Superfamily: NADH dehydrogenase (ubiquinone) chain 6
C:Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;

Query Match 77.3%; Score 34; DB 1; Length 203;
Best Local Similarity 55.6%; Pred. No. 15;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 HVSAALGNLL 9
| | | | |
DB 164 HTSAIGNIM 172

RESULT 2

S02156
NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 6 - Podospora anserina mitochondrion
C:Species: mitochondrion Podospora anserina
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 03-Jun-2002
C:Accession: S02156
R:Cummings, D.J.; Domenico, J.M.
J. Mol. Biol. 204, 815-839, 1988
A:Title: Sequence analysis of mitochondrial DNA from Podospora anserina. Pervasiveness of
A:Reference number: S02153; MUID:89125610; PMID:2975708
A:Accession: S02156
A:Molecule type: DNA
A:Residues: 1-221 <CUM>
A:Cross-References: EMBL:X14486; NID:gi13303; PIDN:CAA32648.1; PID:gi13304
C:Genetics:
A:Genome: mitochondrion
A:Genetic code: SGC3
C:Superfamily: NADH dehydrogenase (ubiquinone) chain 6
C:Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;

Query Match 77.3%; Score 34; DB 1; Length 221;
 Best Local Similarity 44.4%; Pred. No. 17;
 Matches 4; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 HVSAIGNLL 9
 |::|||
 Db 181 HITAGNIM 189

RESULT 3
 T01920
 probable RNA-directed RNA polymerase (EC 2.7.7.48) - Arabidopsis thaliana
 N:Alternate names: protein T2B4.110
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 26-Feb-1999 #sequence_revision 26-Feb-1999 #text_change 10-Dec-1999
 R:Strong, C.; Graves, T.; Duckels, G.
 submitted to the EMBL Data Library, July 1998
 A:Description: The sequence of A. thaliana P2P3.
 A:Reference number: Z14455
 A:Accession: T01920
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1133 <STR>
 A:Cross-references: EMBL:AF080120; NID:G3600045; PID:G3600048
 A:Experimental source: cultivar Columbia
 R:Sevan, M.; Pohl, T.; Weizengger, T.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X.; Schuel
 submitted to the Protein Sequence Database, May 1999
 A:Reference number: Z16098
 A:Accession: T08192
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1133 <BEV>
 A:Cross-references: EMBL:AL049876
 A:Experimental source: cultivar Columbia; BAC clone T22B4
 C:Genetics:
 A:Map position: 4
 A:Introns: 184/3; 820/2; 870/3
 A:Note: F2P3.11; T2B4.110
 C:Superfamily: Arabidopsis probable RNA-directed RNA polymerase
 C:Keywords: nucleotidyltransferase

Query Match 77.3%; Score 34; DB 2; Length 1133;
 Best Local Similarity 66.7%; Pred. No. 11e+02;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 HVSAIGNLL 9
 |::|||
 Db 656 HLSMGNML 664

RESULT 4
 D83079
 hypothetical protein PA4536 [imported] - Pseudomonas aeruginosa (strain PAO1)
 C:Species: Pseudomonas aeruginosa
 C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
 C:Accession: D83079
 R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
 .; Lory, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A:Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic patho
 A:Reference number: A82950; MUID:20437337; PMID:10984043
 A:Accession: D83079
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-149 <SPO>
 A:Cross-references: GB:AE004867; GB:AE004091; NID:G9950769; PIDN:ARG07924.1; GSPDB:GN001
 A:Experimental source: strain PAO1
 C:Genetics:
 A:Gene: PA4536

Query Match 75.0%; Score 33; DB 2; Length 149;

Best Local Similarity 66.7%; Pred. No. 17;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 HVSAIGNLL 9
 |::|||
 Db 51 HIEALGTL 59

RESULT 5
 S28850
 membrane protein PMP22, peroxisomal - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 08-Oct-1999
 C:Accession: S28850
 R:Kaldi, K.; Diestelkoetter, P.; Stenbeck, G.; Auerbach, S.; Jaekle, U.; Maegert, H.J.;
 PBS Lett. 315, 217-222, 1993
 A:Title: Membrane topology of the 22 kDa integral peroxisomal membrane protein.
 A:Reference number: S28850; MUID:93138077; PMID:8422909
 A:Accession: S28850
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-194 <KAL>
 A:Cross-references: EMBL:X70223; NID:G297436; PIDN:CAA49756.1; PID:G297437
 C:Keywords: membrane protein; peroxisome

Query Match 75.0%; Score 33; DB 2; Length 194;
 Best Local Similarity 87.5%; Pred. No. 24;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 VSALGNLL 9
 :|||
 Db 44 LSALGNLL 51

RESULT 6
 T30751
 hypothetical protein 149R - Molluscum contagiosum virus 1
 N:Alternate names: MC149R
 C:Species: Molluscum contagiosum virus 1
 C:Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 16-Feb-2001
 C:Accession: T30751
 R:Senkevich, T.G.; Bugert, J.J.; Sisler, J.R.; Koonin, E.V.; Darai, G.; Moss, B.
 Science 273, 813-816, 1996
 A:Title: Genome sequence of a human tumorigenic poxvirus: Prediction of specific host r
 A:Reference number: Z20876; MUID:96325459; PMID:8670425
 A:Accession: T30751
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-308 <SEN>
 A:Cross-references: EMBL:U60315; NID:G1491943; PIDN:AAC55277.1; PID:G1492092
 C:Genetics:
 A:Note: MC149R
 C:Superfamily: Molluscum contagiosum virus 1 hypothetical protein 149R

Query Match 75.0%; Score 33; DB 2; Length 308;
 Best Local Similarity 77.8%; Pred. No. 40;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 HVSAIGNLL 9
 |::|||
 Db 174 HVSAQGDLL 182

RESULT 7
 S76196
 hypothetical protein - Synechocystis sp. (strain PCC 6803)
 C:Species: Synechocystis sp.
 A:Variety: PCC 6803
 C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 08-Oct-1999
 C:Accession: S76196
 R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.,
 o, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasu
 DNA Res. 3, 109-136, 1996

A;Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* s.
A;Reference number: S74322; MUID:97061201; PMID:8905231
A;Accession: S76196
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-506 <KAN>
A;Cross-references: EMBL:D90914; GB:AB001339; NID:gl653477; PIDN:BAAL8455.1; PID:dl01918
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 75.0%; Score 33; DB 2; Length 506;
Best Local Similarity 75.0%; Pred. No. 70;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 HVSAALGNL 8
|:|||||
Db 309 HLQALGNL 316

RESULT 8
A44165
1-phosphatidylinositol-4,5-bisphosphate phosphodiesterase (EC 3.1.4.11) - slime mold (Di
N;Alternate names: phosphoinositide-specific phospholipase C
C;Species: Dictyostellium discoideum
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 28-Jul-2000
C;Accession: A44165
R;Drayer, A.L.; van Haastert, P.J.
J. Biol. Chem. 267, 18387-18392, 1992
A;Title: Molecular cloning and expression of a phosphoinositide-specific phospholipase C
A;Reference number: A44165
A;Accession: A44165
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-801 <DRA>
A;Cross-references: GB:M95783; NID:gl67846; PIDN:AAA33235.1; PID:gl67847
A;Note: sequence extracted from NCBI backbone (NCBIP:113302)
C;Superfamily: 1-phosphatidylinositol-4,5-bisphosphate phosphodiesterase delta-1; 1-phos
phosphodiesterase domain Y homology; calmodulin repeat homology; pleckstrin repeat homol
C;Keywords: EF hand; phosphoprotein; phosphoric diester hydrolase
F;324-465/Domain: 1-phosphatidylinositol-4,5-bisphosphate phosphodiesterase domain X hom
F;538-656/Domain: 1-phosphatidylinositol-4,5-bisphosphate phosphodiesterase domain Y hom
F;658-769/Domain: protein kinase C C2 region homology <KC2>

Query Match 75.0%; Score 33; DB 1; Length 801;
Best Local Similarity 85.7%; Pred. No. 1.2e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 HVSAALGN 7
|:|||||
Db 470 HVSAVGN 476

RESULT 9
S69745
hypothetical protein YDR360w - yeast (*Saccharomyces cerevisiae*)
C;Species: *Saccharomyces cerevisiae*
C;Date: 23-Aug-1996 #sequence_revision 06-Sep-1996 #text_change 19-Apr-2002
C;Accession: S69745
submitted to the EMBL Data Library, June 1995
A;Description: The sequence of *S. cerevisiae* cosmid 9476.
A;Reference number: S61148
A;Accession: S69745
A;Molecule type: DNA
A;Residues: 1-128 <DUZ>
A;Cross-references: EMBL:U28372; NID:g849170; PID:g2194159; GSPDB:GN00004; MIPS:YDR360w
C;Genetics:
A;Gene: MIPS:YDR360w
A;Cross-references: SGD:S0002768
A;Map position: 4R
C;Superfamily: *Saccharomyces* hypothetical protein YDR360w

Query Match 72.7%; Score 32; DB 2; Length 128;

Best Local Similarity 100.0%; Pred. No. 24;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 SALGNLL 9
|:|||||
Db 67 SALGNLL 73

RESULT 10
D75073
hit-like protein PAB0657 - *Pyrococcus abyssi* (strain Orey)

C;Species: *Pyrococcus abyssi*
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000
C;Accession: D75073
R;anonymous, Genoscope
submitted to the EMBL Data Library, July 1999
A;Description: *Pyrococcus abyssi* genome sequence: insights into archaeal chromosome str
A;Reference number: A75001
A;Accession: D75073
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-147 <KAW>
A;Cross-references: GB:AJ248286; GB:AL096836; NID:g5458366; PIDN:CAB49889.1; PID:g54584
A;Experimental source: strain Orsay
C;Genetics:
A;Gene: PAB0657
C;Superfamily: protein kinase C inhibitor; histidine triad homology
F;3-99/Domain: histidine triad homology <HT>

Query Match 72.7%; Score 32; DB 2; Length 147;
Best Local Similarity 62.5%; Pred. No. 28;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 HVSAALGNL 8
|:|||||
Db 41 HVTISGNL 48

RESULT 11
T17736
probable superoxide dismutase (EC 1.15.1.1) (Cu-Zn) - *Chlorella virus PBCV-1*
C;Species: *Chlorella virus PBCV-1*
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 05-May-2000
C;Accession: T17736
R;Graves, M.V.; Van Etten, J.L.
submitted to the EMBL Data Library, May 1999
A;Reference number: Z18806
A;Accession: T17736
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-187 <GRA>
A;Cross-references: EMBL:U42580; NID:g4028896; PIDN:AAC96613.1
A;Experimental source: specific host *Chlorella* strain NC64A
C;Genetics:
C;Function:
A;Description: catalyzes the dismutation of 2 molecules of peroxide radical to dioxygen
C;Superfamily: superoxide dismutase (Cu-Zn)
C;Keywords: copper; metalloprotein; oxidoreductase; zinc
F;76-78,93,150/Binding site: copper (His) #status predicted
F;87-176/Disulfide bonds: #status predicted
F;93,101,110,113/Binding site: zinc (His, His, Asp) #status predicted
F;173/Active site: Arg #status predicted

Query Match 72.7%; Score 32; DB 2; Length 187;
Best Local Similarity 66.7%; Pred. No. 37;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 HVSAALGNLL 9
|:|||||
Db 110 HVGDIGNLL 118

RESULT 12

T49580
Probable oligomycin sensitivity conferring protein (ATP5) [imported] - Neurospora crassa
N;Alternate names: protein B208.280
C;Species: Neurospora crassa
C;Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000
C;Accession: T49580
R;Schulte, U.; Algn, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,
submitted to the Protein Sequence Database, May 2000
A;Reference number: 225022
A;Accession: T49580
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-220 <SCH>
A;Cross-references: EMBL:AL355930; GSPDB:GN00116; NCSP:B208.280
A;Experimental source: BAC clone B208; strain OR74A
C;Genetics:
A;Gene: NCSP:B208.280
A;Map position: 6
A;Introns: 48/3; 152/3; 183/3

Query Match 72.7%; Score 32; DB 2; Length 220;
Best Local Similarity 75.0%; Pred. No. 44;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 HVSLGNLL 9

:::|||||

Db 65 ISSLGNLL 72

RESULT 13

T35940
Probable transport permease protein - Streptomyces coelicolor
C;Species: Streptomyces coelicolor
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jan-2000
C;Accession: T35940
R;Seeger, K.J.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, January 1999
A;Reference number: 221551
A;Accession: T35940
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-312 <SEE>
A;Cross-references: EMBL:AL035206; PIDN:CA22762.1; GSPDB:GN00070; SCOEDB:SC9B5.20
A;Experimental source: strain A3(2)
C;Genetics:
C;Superfamily: oligopeptide permease protein oppB

Query Match 72.7%; Score 32; DB 2; Length 312;
Best Local Similarity 55.6%; Pred. No. 65;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 HVSLGNLL 9

:::|||||

Db 254 HTASLGNLI 262

RESULT 14

A40708
Basic-helix-loop-helix protein ato - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 16-Feb-1994 #sequence_revision 18-Nov-1994 #text_change 21-Jul-2000
C;Accession: A40708
R;Jarman, A.P.; Grau, Y.; Jan, L.Y.; Jan, Y.N.
Cell 73, 1307-1321, 1993
A;Title: atonal is a proneural gene that directs chordotonal organ formation in the Dros
A;Reference number: A40708; MUID:93313961; PMID:8324823
A;Accession: A40708
A;Status: preliminary
A;Molecule type: nucleic acid
A;Residues: 1-312 <JAR>
A;Cross-references: GB:L36646; NID:g551565; PIDN:AAA21879.1; PID:g551566

A;Note: sequence extracted from NCBI backbone (NCBIN:135094, NCBIP:135095)
C;Genetics:
A;Gene: FlyBase:ato
A;Cross-references: FlyBase:FBgn0010433

Query Match 72.7%; Score 32; DB 2; Length 312;
Best Local Similarity 66.7%; Pred. No. 65;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 HVSLGNLL 9

:::|||||

Db 303 YISALGDL 311

RESULT 15

AE0142
molybdenum cofactor biosynthesis protein A [imported] - Yersinia pestis (strain CO92)
C;Species: Yersinia pestis
C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Nov-2001
C;Accession: AE0142
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.
deno-tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,
Nature 413, 523-527, 2001
A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A;Reference number: AB0001; MUID:21470413; PMID:11586360

A;Accession: AE0142

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-326 <KUR>

A;Cross-references: GB:AL590842; PIDN:CAC90000.1; PID:g15979222; GSPDB:GN00175

C;Genetics:

A;Gene: moaa

C;Superfamily: Escherichia coli molybdopterin biosynthesis protein moaa

Query Match 72.7%; Score 32; DB 2; Length 326;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VSALGNL 8

|||||

Db 262 VSALGNL 268

Search completed: January 20, 2004, 17:03:32
Job time : 11.3435 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 20, 2004, 16:41:41 ; Search time 4.74046 Seconds
(without alignments)
89.283 Million cell updates/sec

Title: US-10-089-549-11
Perfect score: 44
Sequence: 1 HVSAIGNLL 9

Scoring table: BLOSUM62
Gapop: 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	34	77.3	203	1 NU6M_TRIRU	Q36836 trichophyto
2	34	77.3	221	1 NU6M_PODAN	P15959 podospora a
3	33	75.0	193	1 PMP2_MOUSE	P42925 mus musculus
4	33	75.0	193	1 PMP2_RAT	Q07066 rattus norv
5	33	75.0	194	1 LPCA_BUCAP	Q8k9r9 buchnera ap
6	33	75.0	506	1 YF36_SVNY3	P74360 synecocyst
7	33	75.0	801	1 PIPA_DICDI	Q02158 dictyosteli
8	32	72.7	187	1 SODC_CHVPI	Q90023 paramecium
9	32	72.7	220	1 ATPO_NEUCR	Q9p602 neurospora
10	32	72.7	312	1 ATO_DROME	P48987 drosophila
11	32	72.7	326	1 MOAA_VERPE	Q8z9w5 versinia pe
12	32	72.7	888	1 SYA_ZYMMO	Q9rnn8 zymomonas m
13	32	72.7	1180	1 EX58_ECOLI	P08394 escherichia
14	31	70.5	402	1 PA11_MOUSE	P22777 mus musculus
15	31	70.5	410	1 YCAQ_ECOLI	P75843 escherichia
16	31	70.5	1056	1 K125_ARATH	P82266 arabidopsis
17	31	70.5	1842	1 PASZ_SCHPO	Q10289 s fatty aci
18	30	68.2	154	1 SODC_SCHPO	P28758 schizosacch
19	30	68.2	262	1 TRUA_PYRHO	O58941 pyrococcus
20	30	68.2	263	1 TRUA_PYRAB	Q9uz23 pyrococcus
21	30	68.2	266	1 PTRC_KLEPN	P37082 klebsiella
22	30	68.2	411	1 BHB2_MOUSE	Q35185 mus musculus
23	30	68.2	411	1 BHB2_RAT	Q35780 rattus norv
24	30	68.2	412	1 BHB2_HUMAN	O14503 homo sapien
25	30	68.2	435	1 ERF1_TETTH	Q9u8u5 tetrahymena
26	30	68.2	448	1 RMUC_ZYMMO	Q9req3 zymomonas m
27	30	68.2	459	1 PUCF_RHOSH	Q02443 rhodobacter
28	30	68.2	458	1 CWF8_SCHPO	O14011 schizosacch
29	30	68.2	510	1 FOLC_CAEEL	Q09509 caenorhabdi
30	30	68.2	574	1 MUP1_YEAST	P50276 saccharomyc
31	30	68.2	631	1 ADAS_DROME	Q9v778 drosophila
32	30	68.2	631	1 PAB3_HUMAN	Q9h361 homo sapien
33	30	68.2	701	1 TBX2_MOUSE	Q60707 mus musculus

34 30 68.2 702 1 TBX2_HUMAN Q13207 homo sapien
35 30 68.2 707 1 GCVK_HCMVA P16788 human cytom
36 30 68.2 720 1 AGAL_STRMU P27756 streptococc
37 30 68.2 936 1 CAPP_RHOMR Q59757 rhodothermu
38 30 68.2 1148 1 ABA2_HUMAN Q9nti2 homo sapien
39 30 68.2 1148 1 ABA2_MOUSE P98200 mus musculus
40 30 68.2 1297 1 PUR4_PASMU Q9c1w4 pasteurella
41 30 68.2 2200 1 LAR_CAEEL Q9bmm8 caenorhabdi
42 29 65.9 158 1 SODC_BRUPA P41962 brugia paha
43 29 65.9 166 1 SODC_CARCR P80174 caretta car
44 29 65.9 167 1 OB_SMICR Q9xaw9 sminthopsis
45 29 65.9 194 1 PMP2_HUMAN Q9nr77 homo sapien

ALIGNMENTS

RESULT 1
NU6M_TRIRU STANDARD; PRT; 203 AA.
ID Q36836;
AC Q36836;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE NADH-ubiquinone oxidoreductase chain 6 (EC 1.6.5.3).
OS ND6 OR NADH6.
GN Trichophyton rubrum.
OG Mitochondrion.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Onygenales; Arthrodermataceae; mitosporic Arthrodermataceae;
OC Trichophyton.
OX NCBI TaxID=5551;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-IP 1817.89; PubMed=8593686;
RX MEDLINE=96132111; PubMed=8593686;

RA de Bievre C., DuJon B.;
RT "Organisation of the mitochondrial genome of Trichophyton rubrum. DNA sequence analysis of the ND4 gene, the ATPase subunit-6 gene, the ribosomal RNA small-subunit gene, the ND6 gene, the COXIII gene, the ATPase subunit-8 gene and six tRNA genes that correspond respectively to the tyrosine, lysine, glutamine, asparagine, isoleucine and tryptophan isoacceptors.";
RL Curr. Genet. 28:553-559(1995).
CC -!- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X88896; CAA61356.1; -;
CC PIR; S65033; S65033.
CC InterPro; IPR001457; Oxidored_q3.
CC Pfam; PF00499; Oxidored_q3; 1.
CC Oxidoreductase; NAD; Ubiquinone; Mitochondrion.
CC KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion.
CC SQ SEQUENCE 203 AA; 22656 MW; 046DAB3716E02802 CRC64;
CC -----

Query Match 77.3%; Score 34; DB 1; Length 203;
Best Local Similarity 55.6%; Pred. No. 7.8;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 HVSAIGNLL 9
|||:::
Db 164 HTSAIGNIM 172

RESULT 2
NU6M_PODAN STANDARD; PRT; 221 AA.
ID NU6M_PODAN

Db 43 LSLAGNLL 50
:|||||

RESULT 4
PMP2_RAT STANDARD; PRT; 193 AA.
AC Q07066;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Peroxisomal membrane protein 2 (22 kDa peroxisomal membrane protein).
GN PMP2 OR PMP22.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
SEQUENCE FROM N.A., AND SEQUENCE OF 1-18; 30-36; 134-138 AND 144-153.
RP TISSUE=Liver;
RX MEDLINE=93138077; PubMed=8422909;
RA Kaldi K., Diestelkoetter P., Stenbeck G., Auerbach S., Jaekle U.,
RA Maegert H.-J., Wieland F.T., Just W.W.;
RT "Membrane topology of the 22 kDa integral peroxisomal membrane
protein.";
RL FEBS Lett. 315:217-222(1993).
CC -!- FUNCTION: Seems to be involved in pore forming activity and may
CC contribute to the unspecific permeability of the peroxisomal
CC membrane.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Peroxisomal.
CC -!- SIMILARITY: Belongs to the peroxisomal membrane protein PMP2/4
CC family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X70223; CAA49756.1; -.
CC PIR; S28850; S28850.
CC Pfam; PF04117; Mpv17_PMP22; 1.
CC Transmembrane; Peroxisome.
CC INIT_MET 0 0
FT DOMAIN 1 29 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 30 50 POTENTIAL.
FT DOMAIN 51 73 PEROXISOMAL (POTENTIAL).
FT TRANSMEM 74 94 POTENTIAL.
FT DOMAIN 95 112 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 113 133 POTENTIAL.
FT DOMAIN 134 171 PEROXISOMAL (POTENTIAL).
FT TRANSMEM 172 192 POTENTIAL.
SQ SEQUENCE 193 AA; 22446 MW; 14ED1D072852455C CRC64;

Query Match 75.0%; Score 33; DB 1; Length 193;
Best Local Similarity 87.5%; Pred. NO. 12;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 VSLAGNLL 9
:|||||
Db 43 LSLAGNLL 50
:
RESULT 5
LPCA_BUCAP STANDARD; PRT; 194 AA.
AC Q8K9R9;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Phosphoheptose isomerase (EC 5.-.-.-).

GN LPCA OR GMHA OR BUSG241.
OS Buchnera aphidicola (subsp. Schizaphis graminum).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Buchnera.
OX NCBI_TaxID=98794;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=22084549; PubMed=12089438;
RA Tamas I., Klason L., Canbaeck B., Naeslund A.K., Eriksson A.-S.,
RA Wernegreen J.J., Sandstroem J.P., Moran N.A., Andersson S.G.E.;
RT "50 million years of genomic stasis in endosymbiotic bacteria.";
RL Science 296:2376-2379(2002).
CC -!- FUNCTION: Involved in synthesis of glyceromannoheptose 7-phosphate
CC (By similarity).
CC -!- PATHWAY: Inner core lipopolysaccharide biosynthesis.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: BELONGS TO THE SIS FAMILY. LPCA SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE014100; AAM67800.1; -.
CC HAMAP; MF 00067; -; 1.
CC InterPro; IPR004515; GmbA.
CC InterPro; IPR001347; SIS.
CC Pfam; PF01380; SIS; 1.
CC TIGRFAMs; TIGR00441; gmbA; 1.
CC Isomerase; Lipopolysaccharide biosynthesis; Complete proteome.
SQ SEQUENCE 194 AA; 21631 MW; 803DA68727B36830 CRC64;
Query Match 75.0%; Score 33; DB 1; Length 194;
Best Local Similarity 71.4%; Pred. NO. 12;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 HVYALGN 7
:|||||
Db 87 HISAIGN 93
:
RESULT 6
YF26_SYNY3 STANDARD; PRT; 506 AA.
AC P74360;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein sll1526.
GN SLL1526.
OS Synecocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synecocystis.
OX NCBI_TaxID=1148;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=97061201; PubMed=8905231;
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hiroseawa M., Sugiyama M., Sasamoto S., Kimura T.,
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
RA Shimpo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
RA Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synecocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions.";
RL DNA Res. 3:109-136(1996).
CC -!- SIMILARITY: TO R.PROWAZEKII RP789, RP027 AND RP028.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch.)

CC -----
CC EMBL; D90914; BAA18455.1; --
CC PIR; S76196; S76196.
CC InterPro; IPR000051; SAM bind.
CC Hypothetical protein; Complete proteome.
KW SEQUENCE 506 AA; 57727 MW; B12F35E40BE97C1 CRC64;

Query Match 75.0%; Score 33; DB 1; Length 506;
Best Local Similarity 75.0%; Pred. No. 32;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 HVSAALGNL 8
|: |||||
Db 309 HLQALGNL 316

RESULT 7
PIPA_DICDI STANDARD; PRT; 801 AA.
AC Q02158;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 1-phosphatidylinositol-4,5-bisphosphate phosphodiesterase
DE (EC 3.1.4.11) (PLC) (Phosphoinositide phospholipase C).
GN PIPA OR PLC.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
OX NCBI_TaxID=44689;
[1]
SEQUENCE FROM N.A.

CC STRAIN=NC-4;
CC MEDLINE=92406741; PubMed=1326523;
CC Drayer A.L., van Haastert P.J.;
CC "Molecular cloning and expression of a phosphoinositide-specific
CC phospholipase C of Dictyostelium discoideum.";
CC J. Biol. Chem. 267:18387-18392(1992).
CC -1- FUNCTION: THE PRODUCTION OF THE SECOND MESSENGER MOLECULES
CC DIACYLGLYCEROL (DAG) AND INOSITOL 1,4,5-TRISPHOSPHATE (IP3) IS
CC MEDIATED BY ACTIVATED PHOSPHATIDYLINOSITOL-SPECIFIC PHOSPHOLIPASE
CC C ENZYMES.
CC -1- CATALYTIC ACTIVITY: 1-phosphatidyl-1D-myo-inositol 4,5-
CC bisphosphate + H(2)O = D-myo-inositol 1,4,5-trisphosphate +
CC diacylglycerol.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED AT ALL STAGES; INCREASE IN
CC EXPRESSION IN THE CULMINATING FRUITING BODY AND DURING
CC STARVATION.
CC -1- SIMILARITY: DOMAINS X AND Y ARE CONSERVED IN DIFFERENT FORMS OF
CC PLC AND ARE ESSENTIAL FOR CATALYTIC ACTIVITY.
CC -1- SIMILARITY: Contains 1 C2 domain.
CC -1- SIMILARITY: Contains 1 EF-hand calcium-binding domain.
CC -1- CAUTION: IT IS UNCERTAIN WHETHER MET-1 IS THE INITIATOR.

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch.)

CC -----
CC EMBL; M95783; AAA33235.1; --
CC PIR; A44165; A44165.
CC HSSP; P10688; IDJX.
CC DictyDb; DD05034; pipa.
CC InterPro; IPR000008; C2.
CC InterPro; IPR002048; EF-hand.
CC InterPro; IPR001192; PI PLC.
CC InterPro; IPR000909; PI_PLC_Xdom.

DR InterPro; IPR001711; PI_PLC_Y.
DR Pfam; PF00168; C2; 1.
DR Pfam; PF00388; PI-PLC-X; 1.
DR Pfam; PF00387; PI-PLC-Y; 1.
DR PRINTS; PR00390; PHEPLIPASEC.
DR ProDom; PD001202; PI_PLC_Y; 1.
DR SMART; SM00239; C2; 1.
DR SMART; SM00148; PLCXC; 1.
DR SMART; SM00149; PLCYC; 1.
DR PROSITE; PS00018; EF_HAND; 1.
DR PROSITE; PS00044; C2_DOMAIN_2; 1.
DR PROSITE; PS00007; PIPLC_X_DOMAIN; 1.
DR PROSITE; PS00008; PIPLC_Y_DOMAIN; 1.
KW Hydrolase; Lipid degradation; Transducer; Calcium-binding;
KW Phosphorylation.
FT DOMAIN 322 464 DOMAIN X.
FT DOMAIN 542 652 DOMAIN Y.
FT DOMAIN 661 765 C2 DOMAIN.
FT CA_BIND 490 501 EF-HAND (POTENTIAL).
FT ACT_SITE 337 337 BY SIMILARITY.
FT ACT_SITE 382 382 BY SIMILARITY.
FT MOD_RES 524 524 PHOSPHORYLATION (POTENTIAL).
FT MOD_RES 531 531 PHOSPHORYLATION (BY PKA AND PKG)
FT SEQUENCE 801 AA; 91280 MW; DB4F8C829812DD9 CRC64;

Query Match 75.0%; Score 33; DB 1; Length 801;
Best Local Similarity 85.7%; Pred. No. 51;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 HVSAALGN 7
|: |||||
Db 470 HVSAVGN 476

RESULT 8
SODC_CHVP1 STANDARD; PRT; 187 AA.
AC Q90023;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Superoxide dismutase [Cu-Zn] precursor (EC 1.15.1.1).
GN A245R.
OS Paramaricum burearia chlorella virus 1 (PBCV-1).
OC Viruses; dsDNA viruses, no RNA stage; Phycodnaviridae; Chlorovirus.
OX NCBI_TaxID=10506;
[1]
SEQUENCE FROM N.A.
RX MEDLINE=95133167; PubMed=7831789;
RA Lu Z., Li Y., Zhang Y., Kutish G.F., Rock D.L., Van Etten J.L.;
RA "Analysis of 45 kb of DNA located at the left end of the chlorella
RT virology PBCV-1 genome.";
RL Virology 206:339-352(1995).
CC -1- FUNCTION: Destroys radicals which are normally produced within the
CC cells and which are toxic to biological systems.
CC -1- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).
CC -1- COFACTOR: Binds 1 copper ion and 1 zinc ion per subunit (By
CC similarity).
CC -1- SIMILARITY: BELONGS TO THE CU-ZN SUPEROXIDE DISMUTASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch.)

CC -----
CC EMBL; U42580; AAC96613.1; --
CC PIR; T17736; T17736.
CC HSSP; P00441; 4SOD.
CC InterPro; IPR001424; SOD_CU_ZN.

DR Pfam; PF00080; sodcu; 1.
 DR PRINTS; PR00068; CUZNDISMUTASE.
 DR ProDom; PD000469; SOD_CU_ZN; 1.
 DR PROSITE; PS00087; SOD_CU_ZN_1; 1.
 DR PROSITE; PS00332; SOD_CU_ZN_2; 1.
 KW Antioxidant; Oxidoreductase; Metal-binding; Copper; Zinc; Signal.
 FT SIGNAL 1 21 POTENTIAL.
 FT CHAIN 22 187 SUPEROXIDE DISMUTASE [CU-ZN].
 FT METAL 76 76 COPPER (BY SIMILARITY).
 FT METAL 78 78 COPPER (BY SIMILARITY).
 FT METAL 93 93 COPPER AND ZINC (BY SIMILARITY).
 FT METAL 101 101 ZINC (BY SIMILARITY).
 FT METAL 110 110 ZINC (BY SIMILARITY).
 FT METAL 113 113 ZINC (BY SIMILARITY).
 FT METAL 150 150 COPPER (BY SIMILARITY).
 FT DISULPID 87 176 BY SIMILARITY.
 SQ SEQUENCE 187 AA; 19748 MW; E12366A9274674DD CRC64;
 Query Match 72.7%; Score 32; DB 1; Length 187;
 Best Local Similarity 66.7%; Pred. No. 19;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 1 HVSAALGNLL 9
 DB 110 HVGDLGNLL 118
 RESULT 9
 ID ATPO_NEUCR STANDARD; PRT; 220 AA.
 AC Q9P602;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE ATP synthase oligomycin sensitivity conferral protein, mitochondrial
 DE precursor (EC 3.6.3.14) (OSCP) (ATP synthase chain 5).
 GN ATP-5 OR B208.150.
 OS Neurospora crassa.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
 OX NCBI_TaxID=5141;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=74-OR23-1A / FGSC 987;
 RX PubMed=12655011;
 RA Mannheim J.D., Fortmann B., Nyakatura G., Kempken F., Maier J.,
 RA Hohelsel J.D.,
 RA "What's in the genome of a filamentous fungus? Analysis of the
 RT Neurospora genome sequence."
 RL Nucleic Acids Res. 31:1944-1954(2003).
 CC -1- FUNCTION: This protein seems to be part of the stalk that links
 CC CF(0) to CF(1). It either transmits conformational changes from
 CC CF(0) into CF(1) or is implicated in proton conduction (By
 CC similarity).
 CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (In) = ADP + phosphate +
 CC H(+) (Out).
 CC -1- SUBUNIT: F-type ATPases have 2 components, CF(1) - the catalytic
 CC core - and CF(0) - the membrane proton channel. CF(1) has five
 CC subunits: alpha(3), beta(3), gamma(1), delta(1), epsilon(1). CF(0)
 CC has three main subunits: A, B and C (By similarity).
 CC -1- SUBCELLULAR LOCATION: Mitochondrial.
 CC -1- SIMILARITY: BELONGS TO THE ATPASE DELTA CHAIN FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AL355930; CAB91368.2; -.

DR PIR; T49580; T49580.
 DR InterPro; IPR000711; ATPynt_OSCP.
 DR Pfam; PF00213; OSCP; 1.
 DR PRINTS; PR00125; ATPASEDELTA.
 DR TIGRPFAM; TIGR01145; ATP_synt_delta; 1.
 DR PROSITE; PS00389; ATPASE_DELTA; 1.
 KW Hydrolase; ATP synthesis; CF(1); Hydrogen ion transport;
 KW Mitochondrion; Transit peptide.
 FT TRANSIT 1 220 MITOCHONDRION (POTENTIAL).
 FT CHAIN ? 220 ATP SYNTHASE OLIGOMYCIN SENSITIVITY
 FT CONFERRAL PROTEIN. CRC64;
 SQ SEQUENCE 220 AA; 23047 MW; 7C6E82BD511C2C81 CRC64;
 Query Match 72.7%; Score 32; DB 1; Length 220;
 Best Local Similarity 75.0%; Pred. No. 22;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 2 VSALGNLL 9
 DB 65 ISSLGNLL 72
 RESULT 10
 ID ATO_DROME STANDARD; PRT; 312 AA.
 AC P48987; Q9VHU0;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Atonal protein.
 DE ATO OR CG7508.
 GN Drosophila melanogaster (Fruit fly).
 OS Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A., FUNCTION, AND TISSUE SPECIFICITY.
 RC STRAIN=Oregon-R;
 RX MEDLINE=93313961; PubMed=8324823;
 RA Jarman A.P., Grau Y., Jan L.Y., Jan Y.N.;
 RA "Atonal is a proneural gene that directs chordotonal organ formation
 RT in the Drosophila peripheral nervous system."
 RL Cell 73:1307-1321(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkeley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Abghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Dang Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K.J., Doup L.E., Downes M., Dugan-Cochran S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 RA Hoston D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.B., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry J., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,

RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Klamos I., Simpson M., Stupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zavari J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RA "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 RN [3]
 RP FUNCTION
 RX MEDLINE=94255014; PubMed=8196767;
 RA Jarman A.P., Grell E.H., Ackerman L., Jan L.Y., Jan Y.N.;
 RT "Atonal is the proneural gene for *Drosophila* photoreceptors.";
 RL Nature 369:398-400(1994).
 CC -1- FUNCTION: DEVELOPMENTAL PROTEIN INVOLVED IN NEUROGENESIS. REQUIRED
 CC FOR THE FORMATION OF CHORDONAL ORGANS AND PHOTORECEPTORS. SEEMS
 CC TO BIND TO E BOXES. SPECIFICALLY REQUIRED FOR THE PHOTORECEPTOR R8
 CC SELECTION.
 CC -1- SUBUNIT: EFFICIENT DNA BINDING REQUIRES DIMERIZATION WITH ANOTHER
 CC BHLH PROTEIN. FORMS HETERODIMER WITH DAUGHTERLESS.
 CC -1- SUBCELLULAR LOCALIZATION: Nuclear.
 CC -1- TISSUE SPECIFICITY: PRONEURAL CLUSTERS AND SENSE ORGAN PRECURSORS
 CC OF THE CHORDONAL ORGANS, OPTIC FURROW OF THE EYE-ANTENNAL DISK
 CC AND DEVELOPING BRAIN LOBE.
 CC -1- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
 CC TRANSCRIPTION FACTORS. "ATONAL" SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; L36646; AAA21879.1; -;
 CC EMBL; AE003678; AAF54209.1; -;
 CC PIR; A40708; A40708.
 CC FlyBase; FBgn010433; ato.
 CC GO; GO:0007420; P:brain development; IMP.
 CC GO; GO:0007173; P:EGF receptor signaling pathway; IGI.
 CC GO; GO:0007605; P:hearing; IMP.
 CC GO; GO:0007422; P:peripheral nervous system development; NAS.
 CC GO; GO:0045464; P:R8 cell fate specification; NAS.
 CC GO; GO:0007224; P:smoothed receptor signaling pathway; IGI.
 CC InterPro; IPR001092; HLH_basic.
 CC Pfam; PF00010; HLH; 1.
 CC SMART; SM00353; HLH; 1.
 CC PROSITE; PS00038; HLH 1; FALSE_NEG.
 CC PROSITE; PS00888; HLH 2; 1.
 CC Neurogenesis; Differentiation; Developmental protein; Nuclear protein;
 KW Transcription regulation; DNA-binding.
 FT DNA_BIND 255 267 BASIC DOMAIN
 FT DOMAIN 268 308 HELIX-LOOP-HELIX MOTIF (BY SIMILARITY).
 FT CONFLICT 149 149 G -> A (IN REF. 1).
 SQ SEQUENCE 312 AA; 34116 MW; 069479287438F456 CRC64;
 Query Match 72.7%; Score 32; DB 1; Length 312;
 Best Local Similarity 66.7%; Pred. No. 32;
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 QY 1 HVSAAGNLL 9
 Db 303 YISALGDL 311

RESULT 11
 MOAA_YERPE

ID MOAA_YERPE STANDARD; PRT; 326 AA.
 AC Q8ZGW5;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Molybdenum cofactor biosynthesis protein A.
 GN MOAA OR YP01159 OR Y3023.
 OS *Yersinia pestis*.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Yersinia.
 OX NCBI_TaxID=632;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=CO-92 / Biovar Orientalis;
 RX MEDLINE=21470413; PubMed=11586360;
 RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
 RA Prentice M.B., Sebahia M., James K.D., Churcher C., Mungall K.L.,
 RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdano-Tarraga A.M.,
 RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
 RA Felwell T., Hamlin N., Holroyd S., Jagsels K., Karlyshev A.V.,
 RA Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,
 RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.;
 RT "Genome sequence of *Yersinia pestis*, the causative agent of plague.";
 RL Nature 413:523-527(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX STRAIN=KIM5 / Biovar Mediaevalis;
 RX MEDLINE=22137863; PubMed=12142430;
 RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
 RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
 RA Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
 RA Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,
 RA Perry R.D.;
 RT "Genome sequence of *Yersinia pestis* KIM.";
 RL J. Bacteriol. 184:4601-4611(2002).
 CC -1- FUNCTION: Together with moac, is involved in the conversion of a
 CC guanosine derivative (GXP) into molybdopterin precursor Z (By
 CC similarity).
 CC -1- COFACTOR: Binds 1 3Fe-4S cluster (By similarity).
 CC -1- PATHWAY: Molybdenum cofactor biosynthesis; first step.
 CC -1- SIMILARITY: BELONGS TO THE MOAA / NIFB / PQE FAMILY. MOAA
 CC SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AJ141446; CAC90000.1; -;
 CC EMBL; AE013903; AAM86573.1; ALT_INIT.
 CC PIR; AE0142; AE0142.
 CC HAMAP; MF_01225; -; 1.
 CC InterPro; IPR006638; ELP3.
 CC InterPro; IPR000385; MoaA_NifB_PqgE.
 CC Pfam; PF04055; Radical_SAM; 1.
 CC SMART; SM00729; ELP3; 1.
 CC PROSITE; PS01305; MOAA_NIFB_PQEE; 1.
 CC Molybdenum cofactor biosynthesis; Metal-binding; Iron; Iron-sulfur;
 KW 3Fe-4S; Complete proteome.
 FT METAL 23 23 IRON-SULFUR (3FE-4S) (POTENTIAL).
 FT METAL 30 30 IRON-SULFUR (3FE-4S) (POTENTIAL).
 FT METAL 254 254 IRON-SULFUR (3FE-4S) (POTENTIAL).
 FT METAL 271 271 IRON-SULFUR (3FE-4S) (POTENTIAL).
 SQ SEQUENCE 326 AA; 36846 MW; 0F614E339D50A8B5 CRC64;
 Query Match 72.7%; Score 32; DB 1; Length 326;
 Best Local Similarity 100.0%; Pred. No. 33;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 VSALGNL 8


```

Db      262 VSALGNL 268
|||||
RESULT 12
SYA_ZYMO
ID SYA_ZYMO STANDARD; PRT; 888 AA.
AC Q9NN8;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Alanyl-tRNA synthetase (EC 6.1.1.7) (Alanine--tRNA ligase) (AlaRS).
GN ALAS.
OS Zymomonas mobilis.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Sphingomonadales;
OC Sphingomonadaceae; Zymomonas.
OX NCBI_TaxID=542;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 31821 / ZM4 / CP4;
RA Um H.W., Kang H.S.;
RL Submitted (Aug-1999) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: ATP + L-alanine + tRNA(Ala) = AMP +
CC diphosphate + L-alanyl-tRNA(Ala).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF179611; AAD53924.1; -.
DR HAMAP; MF_00036; -.
DR InterPro; IPR003156; DHHAL.
DR InterPro; IPR002318; tRNA-synt_2c.
DR InterPro; IPR006193; tRNA_synt_Ala.
DR Pfam; PF02272; DHHAL; 1.
DR Pfam; PF01411; tRNA-synt_2c; 1.
DR PRINTS; PRO0980; TRNASYNTHAL.
DR TIGRFAMs; TIGR00344; alas; 1.
DR PROSITE; PS50860; AA tRNA LIGASE II ALA; 1.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.
SQ SEQUENCE 888 AA; 97015 MW; 119A8867550A38CD CRC64;

Query Match 72.7%; Score 32; DB 1; Length 888;
Best Local Similarity 75.0%; Pred. No. 92;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 HVSAIGNL 8
|||||
Db 680 HVSAIGNL 687

RESULT 13
EX5B_ECOLI
ID EX5B_ECOLI STANDARD; PRT; 1180 AA.
AC P08394;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Exodeoxyribonuclease V beta chain (EC 3.1.11.5) (Exodeoxyribonuclease
DE V 135 kDa polypeptide).
GN RECB OR R08A OR B280.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87066729; PubMed=3537960;
RA Finch P.W., Storey A., Chapman K.E., Brown K., Hickson I.D.,
RA Emmerson P.T.;
RT "Complete nucleotide sequence of the Escherichia coli recB gene.";
RL Nucleic Acids Res. 14:8573-8582(1986).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=V1000;
RX MEDLINE=20229837; PubMed=10766864;
RA Arnold D.A., Kowalczykowski S.C.;
RT "Facilitated loading of RecA protein is essential to recombination by
RT RecBCD enzyme.";
RL J. Biol. Chem. 275:12261-12265(2000).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [4]
RP SEQUENCE OF 1-11 FROM N.A.
RX MEDLINE=87040734; PubMed=3534791;
RA Finch P.W., Wilson R.E., Brown K., Hickson I.D., Emmerson P.T.;
RT "Complete nucleotide sequence of the Escherichia coli ptr gene
RT encoding protease III.";
RL Nucleic Acids Res. 14:7695-7703(1986).
RN [5]
RP SEQUENCE OF 1093-1180 FROM N.A.
RX MEDLINE=87066730; PubMed=3537961;
RA Finch P.W., Storey A., Brown K., Hickson I.D., Emmerson P.T.;
RT "Complete nucleotide sequence of recD, the structural gene for the
RT alpha subunit of Exonuclease V of Escherichia coli.";
RL Nucleic Acids Res. 14:8583-8594(1986).
CC -1- FUNCTION: REQUIRED FOR EFFICIENT DNA REPAIR; IT CATALYZES THE
CC UNWINDING OF DOUBLE-STRANDED DNA AND THE CLEAVAGE OF SINGLE-
CC STRANDED DNA AND IT STIMULATES LOCAL GENETIC RECOMBINATION.
CC ALL OF THESE ACTIVITIES REQUIRE CONCOMITANT HYDROLYSIS OF ATP.
CC -1- CATALYTIC ACTIVITY: Exonucleolytic cleavage (in the presence of
CC ATP) in either 5'- to 3'- or 3'- to 5'-direction to yield 5'-
CC phosphooligonucleotides.
CC -1- SUBUNIT: CONSIST OF THREE SUBUNITS; RECB, RECC AND RECD.
CC -1- SIMILARITY: BELONGS TO THE HELICASE FAMILY. UVRD SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X04581; CAA28250.1; -.
DR EMBL; AF179304; AAD56369.1; -.
DR EMBL; U29581; AAB40467.1; -.
DR EMBL; AE000365; AAC75859.1; -.
DR EMBL; X06227; CAA29577.1; -.
DR EMBL; X04582; CAA28252.1; -.
DR PIR; A25532; NCECK5.
DR HSSP; P56255; 1PUR.
DR Ecogene; EG10824; recB.
DR InterPro; IPR004586; RecB.
DR InterPro; IPR000212; UvrD-helicase.
DR Pfam; PF00580; UvrD-helicase; 1.
DR TIGRFAMs; TIGR00609; recB; 1.
KW Hydrolase; Nuclease; Exonuclease; Endonuclease; Helicase; ATP-binding;
KW DNA repair; Complete proteome.
NP BIND 23 30 ATP.
SQ SEQUENCE 1180 AA; 133958 MW; F9AC331808B8F281 CRC64;

```

Query Match 72.7%; Score 32; DB 1; Length 1180;
Best Local Similarity 77.8%; Pred. No. 1-2e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 HVSAALGNLL 9
| | | | |
Db 834 HQSALGRLL 842

RESULT 14
PAI1_MOUSE
ID PAI1_MOUSE STANDARD; PRT; 402 AA.
AC P2277;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Plasminogen activator inhibitor-1 precursor (PAI-1) (Endothelial
DE plasminogen activator inhibitor) (PAI).
GN SERPINE1 OR PAI1 OR PLANHI OR MR1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90158593; PubMed=2406566;
RA Prendergast G.C.; Diamond L.B.; Dahl D.; Cole M.D.;
RT "The c-myc-regulated gene mrl encodes plasminogen activator inhibitor
1.";
RL Mol. Cell. Biol. 10:1265-1269(1990).
CC -1- FUNCTION: THIS INHIBITOR ACTS AS "BAIT" FOR TISSUE PLASMINOGEN
CC ACTIVATOR, UROKINASE, AND PROTEIN C. ITS RAPID INTERACTION WITH
CC TPA MAY FUNCTION AS A MAJOR CONTROL POINT IN THE REGULATION OF
CC FIBRINOLYSIS.
CC -1- MISCELLANEOUS: PAI1 IS INACTIVATED BY PROTEOLYTIC ATTACK OF THE
CC UROKINASE-TYPE (U-PA) AND THE TISSUE-TYPE (TPA), CLEAVING THE
CC 369(R)-370(M) BOND.
CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M33960; AAA39887.1; -.
CC PIR; A34761; A34761.
CC HSSP; P05121; 1A7C.
CC MGD; MGI:978608; Serpinel.
CC InterPro; IPR000215; Serpin.
CC Pfam; PF00079; serpin; 1.
CC SMART; SM00093; SERPIN; 1.
CC PROSITE; PS00284; SERPIN; 1.
CC Serpin; Serine protease inhibitor; Plasma; Plasminogen activation;
KW Glycoprotein; Signal.
FT SIGNAL 1 23
FT CHAIN 24 402
FT ACT SITE 369 370
FT CARBOHYD 232 232
FT CARBOHYD 288 288
FT CARBOHYD 352 352
FT SEQUENCE 402 AA; 45170 MW; 765FF1659C70F68C CRC64;
SQ

Query Match 70.5%; Score 31; DB 1; Length 402;
Best Local Similarity 66.7%; Pred. No. 67;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 HVSAALGNLL 9
| | | | |
Db 269 HLSALTNIL 277

RESULT 15
YCAQ_ECOLI STANDARD; PRT; 410 AA.
AC P75843;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein ycaQ.
GN YCAQ OR B0916.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R.; Plunkett G. III; Bloch C.A.; Perna N.T.; Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RN Science 277:1453-1474(1997).
RP SEQUENCE FROM N.A.
RX STRAIN=K12;
RX MEDLINE=97061202; PubMed=8905232;
RA Oshima T., Aiba H., Baba T., Fujita K., Hayashi K., Honjo A.,
RA Ikenoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
RA Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,
RA Sampei G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
RA Yano M., Horiiuchi T.;
RT "A 718-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 12.7-28.0 min region on the linkage map.";
RN DNA Res. 3:137-155(1996).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB000193; AAC74002.1; -.
CC EMBL; D90730; BAA35662.1; -.
CC PIR; C64831; C64831.
CC EcoGene; EG13701; ycaQ.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 410 AA; 47655 MW; 76A8BEE084D7307D CRC64;

Query Match 70.5%; Score 31; DB 1; Length 410;
Best Local Similarity 75.0%; Pred. No. 68;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 HVSAALGNL 8
| | | | |
Db 255 HVEKLGSL 262

Search completed: January 20, 2004, 16:54:15
Job time : 5.74046 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 20, 2004, 16:42:36 ; Search time 25.9695 Seconds
(without alignments)
89.431 Million cell updates/sec

Title: US-10-089-549-11
Perfect score: 44
Sequence: 1 HVSAIAGNLL 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_23:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertibrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	37	84.1	330 2	Q8GKR1 corynebacte
2	37	84.1	330 16	Q8NMB2 corynebacte
3	37	84.1	428 16	Q8F7J7 leptospira
4	35	79.5	214 8	Q8HG50 verticilli
5	35	79.5	250 8	Q8AK64 trichoderma
6	35	79.5	256 8	Q8EHQ5 trichoderma
7	35	79.5	1288 17	Q96XM6 sulfolobus
8	34	77.3	109 16	Q8NT07 corynebacte
9	34	77.3	109 17	Q8TNP4 methanosarc
10	34	77.3	1133 10	Q82504 arabidopsis
11	33	75.0	149 16	Q9HVP1 pseudomonas
12	33	75.0	255 5	Q8ITF4 giardia lam
13	33	75.0	308 12	Q98315 molluscum c
14	33	75.0	559 11	Q8CHX3 mus musculu
15	33	75.0	623 17	Q8PXC9 methanosarc
16	33	75.0	1227 2	Q8L122 pseudomonas

17	33	75.0	1284	2	Q8RQU9	Q8rqu9 bacillus gl
18	32	72.7	128	3	P87286	P87286 saccharomyc
19	32	72.7	130	16	Q8E520	Q8e520 streptococc
20	32	72.7	130	16	Q8DZF4	Q8dzf4 streptococc
21	32	72.7	147	17	Q9V017	Q9v017 pyrococcus
22	32	72.7	312	16	Q9ZBG0	Q9zbg0 streptomyce
23	32	72.7	427	5	Q8T9G2	Q8t9g2 drosophila
24	32	72.7	427	5	Q9VIU4	Q9viu4 drosophila
25	32	72.7	488	10	Q94GE0	Q94ge0 oryza sativ
26	32	72.7	554	16	O07411	O07411 mycobacteri
27	32	72.7	885	2	Q9APM7	Q9apm7 myxococcus
28	32	72.7	888	2	Q9RNN8	Q9rnn8 zymomonas m
29	32	72.7	1180	16	Q8X6M9	Q8x6m9 escherichia
30	32	72.7	1181	16	Q8ZMB6	Q8zmb6 salmonella
31	32	72.7	1181	16	Q8Z419	Q8z419 salmonella
32	32	72.7	1183	16	Q8FEB3	Q8feb3 escherichia
33	32	72.7	1732	11	Q60858	Q60858 mus musculu
34	32	72.7	1976	11	Q63541	Q63541 rattus norv
35	32	72.7	1978	11	Q9WTU3	Q9wtu3 mus musculu
36	32	72.7	1978	11	O88420	O88420 rattus norv
37	32	72.7	1980	4	Q95788	Q95788 homo sapien
38	32	72.7	1980	4	Q9UQD0	Q9uqd0 homo sapien
39	32	72.7	1980	4	Q9NYX2	Q9nyx2 homo sapien
40	32	72.7	1988	11	O88421	O88421 rattus norv
41	32	72.7	2147	5	Q8IIH3	Q8iih3 plasmodium
42	32	72.7	4600	5	Q8IFF9	Q8iff9 trypanosoma
43	31.5	71.6	386	17	O8PTW6	O8ptw6 methanosarc
44	31	70.5	98	9	O48436	O48436 bacterioph
45	31	70.5	174	16	Q97RS0	Q97rs0 streptococc

ALIGNMENTS

RESULT 1

Q8GKR1 ID Q8GKR1 PRELIMINARY; PRT; 330 AA.
AC Q8GKR1;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Coproporphyrinogen III oxidase.
GN HEMN.
OS Corynebacterium glutamicum (Brevibacterium flavum).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.
OX NCBI_taxid=17118;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=R;
RA Vertes A.A., Kos P.B., Inui M., Yukawa H.;
RT "Evolutionary Considerations on the Heme Biosynthetic Pathway of
Corynebacteria and Mycobacteria.";
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY150857; AAN72321.1; -
SQ SEQUENCE 330 AA; 35736 MW; 7A60329491FA46C6 CRC64;

Query Match 84.1%; Score 37; DB 2; Length 330;
Best Local Similarity 87.5%; Pred. No. 18;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 HVSAIAGNLL 8

DB 300 HVNAGNLL 307

RESULT 2

Q8NMB2 ID Q8NMB2 PRELIMINARY; PRT; 330 AA.
AC Q8NMB2;
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)

```

DE Coptoporphyrinogen III oxidase and related FeS oxidoreductases (EC
DE 1.3.3.3).
GN CGL2292.
OS Corynebacterium glutamicum (Brevibacterium flavum).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=1718;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RA Nakagawa S.;
RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AP005281; BAB9685.1; -.
DR InterPro: IPR004559; HemN_rel.
DR TIGRPFAM: TIGR00539; hemN_rel. 1.
KW Oxidoreductase; Complete proteome.
SQ SEQUENCE 330 AA; 35704 MW; DAA4E62DA63900CB CRC64;

Query Match 84.1%; Score 37; DB 16; Length 330;
Best Local Similarity 87.5%; Pred. No. 18;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 HVSAAGNL 8
DB 300 HVNAGNL 307

RESULT 3
Q8F7J7 ID Q8F7J7 PRELIMINARY; PRT; 428 AA.
AC Q8F7J7
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical zinc protease (EC 3.4.99.-).
GN LA0948.
OS Leptospira interrogans.
OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
OX NCBI_TaxID=173;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
RA Ren S.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB011278; AAN48147.1; -.
KW Hypothetical protein; Hydrolase; Protease; Complete proteome.
SQ SEQUENCE 428 AA; 48510 MW; 58D0EA80EAC838F1 CRC64;

Query Match 84.1%; Score 37; DB 16; Length 428;
Best Local Similarity 87.5%; Pred. No. 23;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 HVSAAGNL 8
DB 407 HLSAGNL 414

RESULT 4
Q8HG50 ID Q8HG50 PRELIMINARY; PRT; 214 AA.
AC Q8HG50;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE NADH dehydrogenase subunit 6.
GN ND6.
OS Verticillium lecanii.
OC Mitochondrion.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Hypocreales; Mitosporic Hypocreales; Verticillium.
OX NCBI_TaxID=40604;
RN (1)

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=C42;
RA Kouvelis V.N., Typas M.A.;
RT "Structural organization of the mitochondrial genome of the
entomopathogenic fungus Verticillium lecanii: DNA sequence analysis of
the small rRNA subunit, NADH dehydrogenase subunit 1 and cytochrome
oxidase subunit III.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RN (2)
RP SEQUENCE FROM N.A.
RC STRAIN=C42;
RA Kouvelis V.N., Ghikas D.V., Typas M.A.;
RT "An analysis of the complete mitochondrial genome of the
entomopathogenic fungus Verticillium lecanii.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF487277; AAO14669.1; -.
KW Mitochondrion.
SQ SEQUENCE 214 AA; 24174 MW; 91C315769FB7448B CRC64;

Query Match 79.5%; Score 35; DB 8; Length 214;
Best Local Similarity 55.6%; Pred. No. 29;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 HVSAAGNL 9
DB 174 HISTGNIL 182

RESULT 5
Q8SK64 ID Q8SK64 PRELIMINARY; PRT; 250 AA.
AC Q8SK64;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE NADH dehydrogenase subunit VI.
GN ND6.
OS Trichoderma pseudokoningii.
OC Mitochondrion.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Hypocreales; mitosporic Hypocreales; Trichoderma.
OX NCBI_TaxID=51486;
RN (1)
RP SEQUENCE FROM N.A.
RC Wang T.H., Wu Z.H., Liu S.L., Huang W.;
RT "Cloning and sequence analysis of a mitochondrial gene cluster
encoding cytochrome oxidase subunit III, NADH dehydrogenase subunit VI
and nine tRNA molecules from Trichoderma pseudokoningii.";
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY072697; AAL76100.1; -.
DR InterPro: IPR001457; Oxidored_q3.
DR Pfam: PF00499; Oxidored_q3; 1.
KW Mitochondrion.
SQ SEQUENCE 250 AA; 28009 MW; A117E7EF256C2874 CRC64;

Query Match 79.5%; Score 35; DB 8; Length 250;
Best Local Similarity 55.6%; Pred. No. 35;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 HVSAAGNL 9
DB 183 HISTGNIL 191

RESULT 6
Q8SHQ5 ID Q8SHQ5 PRELIMINARY; PRT; 256 AA.
AC Q8SHQ5;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE NADH-ubiquinone oxidoreductase chain 6.
GN ND6.

```

OS Trichoderma reesei (Hypocrea jecorina).
 OG Mitochondrion.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Hypocreales; Hypocreaceae; Hypocrea.
 OX NCBI_TaxID=51453;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21950703; PubMed=11825887;
 RA Chamberg F.S., Bonaccorsi E.D., Ferreira A.J.S., Ramos A.S.P.,
 RA Ferreira Junior J.R., Abrahao-Neto J., Farah J.P.S., El-Dorri H.,
 RA "Elucidation of the Metabolic Fate of Glucose in the Filamentous
 RT Fungus Trichoderma reesei Using Expressed Sequence Tag (EST) Analysis
 RT and cDNA Microarrays";
 RL J. Biol. Chem. 277:13983-13988(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Chamberg F.S., Bonaccorsi E.D., Ferreira A.J.S., Ramos A.S.P.,
 RA Ferreira Junior J.R., Abrahao-Neto J., Farah J.P.S., El-Dorri H.,
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF447590; AAL74172.1; -
 DR InterPro: IPR001457; Oxidored q3.
 DR Pfam: PF00499; oxidored q3; 1.
 KW Ubiquinone; Mitochondrion.
 SQ SEQUENCE 256 AA; 28744 MW; 55EPB10CE0E05CC8 CRC64;
 Query Match 79.5%; Score 35; DB 8; Length 256;
 Best Local Similarity 55.6%; Pred. No. 36;
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 QY 1 HVSAQNLL 9
 Db 189 HISTIGNIL 197
 RESULT 7
 Q96XM6 PRELIMINARY; PRT; 1288 AA.
 AC Q96XM6;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Putative pseudomonapepsin.
 GN ST2492.
 OS Sulfolobus tokodaii.
 OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
 OC Sulfolobus.
 OX NCBI_TaxID=111955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=JCM 10545 / 7;
 RX PubMed=11572479;
 RA Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
 RA Sekine M., Baba S.-I., Aikai A., Kosugi H., Hosoyama A., Fukui S.,
 RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
 RA Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,
 RA Aoki K.-I., Maeda S., Yanagii M., Nishimura M., Yamagishi A.,
 RA Oshina T., Kikuchi H.;
 RT "Complete genome sequence of an aerobic thermoacidophilic
 RT Crenarchaeon, Sulfolobus tokodaii strain7";
 RL DNA Res. 8:123-140(2001).
 DR EMBL; AF000990; BAB67601.1; -
 DR InterPro: IPR000209; Peptidase S8.
 DR Pfam: PF00082; Peptidase S8; 1_
 DR PROSITE; PS00138; SUBTILASE_SER; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 1288 AA; 139443 MW; 54FAC81B90F190AB CRC64;
 Query Match 79.5%; Score 35; DB 17; Length 1288;
 Best Local Similarity 66.7%; Pred. No. 2.1e+02;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 HVSAQNLL 9
 Db 189 HISTIGNIL 197

Db 1239 HISSIGNTL 1247
 RESULT 8
 Q8NT07 PRELIMINARY; PRT; 109 AA.
 ID Q8NT07;
 AC Q8NT07;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Hypothetical protein Cg10509.
 GN Cg10509.
 OS Corynebacterium glutamicum (Brevibacterium flavum).
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacteriineae; Corynebacteriaceae; Corynebacterium.
 OX NCBI_TaxID=1718;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
 RA Nakagawa S.;
 RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032";
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AP005275; BAB97901.1; -
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 109 AA; 11286 MW; 28256A202F3694DB CRC64;
 Query Match 77.3%; Score 34; DB 16; Length 109;
 Best Local Similarity 75.0%; Pred. No. 23;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 HVSAQNLL 8
 Db 26 HISALGSL 33
 RESULT 9
 Q8TNP4 PRELIMINARY; PRT; 675 AA.
 ID Q8TNP4;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Hypothetical protein MA2239.
 GN MA2239.
 OS Methanosarcina acetivorans.
 OC Archaea; Euryarchaeota; Methanococci; Methanosarcinales;
 OC Methanosarcinaceae; Methanosarcina.
 OX NCBI_TaxID=2214;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C2A / ATCC 35395 / DSM 2834;
 RX MEDLINE=21929760; PubMed=11932238;
 RA Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., Macdonald P.,
 RA Fitzhugh W., Calvo S., Engels R., Smirnov S., Ancoor D., Brown A.,
 RA Allen N., Naylor J., Stange-Thomann N., Dearellano K., Johnson R.,
 RA Linton L., McSwan P., McKernan K., Talamas J., Tirrell A., Ye W.,
 RA Zimmer A., Barber R.D., Cann I., Graham D.E., Grahame D.A., Guss A.M.,
 RA Hedderich R., Ingram-Smith C., Kuettnner H.C., Krzycki J.A.,
 RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
 RA Springer T.A., Umayam L.A., White O., White R.H., de Macario E.C.,
 RA Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,
 RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
 RA Metcalf W.W., Birren B.;
 RT "The genome of Methanosarcina acetivorans reveals extensive metabolic
 RT and physiological diversity";
 RL Genome Res. 12:532-542(2002).
 DR EMBL; AF010910; RAM05634.1; -
 DR InterPro: IPR001237; Postsynaptic.
 DR Pfam: PF001440; TPR.
 DR Pfam: PF00515; TPR; 8.
 DR ProDom; PD012428; Postsynaptic; 1.
 DR SMART; SM00028; TPR; 8.
 KW Hypothetical protein; Complete proteome.

SQ SEQUENCE 675 AA; 77117 MW; 551931ADC3EC8680 CRC64;

Query Match 77.3%; Score 34; DB 17; Length 675;
 Best Local Similarity 75.0%; Pred. No. 1.7e+02;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 HVSAALGNL 8
 ||:|||||
 Db 340 HSIILGNL 347

RESULT 10

082504 PRELIMINARY; PRT; 1133 AA.

ID 082504; AC 082504; DT 01-NOV-1998 (TREMBLrel. 08, Created)
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
 DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
 DE F2P3.11 protein (Putative RNA-directed RNA polymerase).
 GN F2P3.11 OR T22B4.110 OR A74G11130.
 OS Arabidopsis thaliana (mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RA WASHU;
 RT "The A. thaliana Genome Sequencing Project."
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RA Strong C., Graves T., Duckels G.;
 RT "The sequence of A. thaliana F2P3."
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RA Waterston R.;
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Bevan M., Pohl T., Weizenegger T., Bancroft I., Mewes H.W.,
 RA Mayer K.F.X., Schueller C.;
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A.
 RA Pohl T., Weizenegger T., Mewes H.W., Lemcke K., Mayer K.F.X.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [7]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF080120; AAC35535.1; -
 DR EMBL; AL049876; CAB43048.1; -
 DR EMBL; AL161331; CAB81214.1; -
 DR Nucleotidyltransferase; RNA-directed RNA polymerase; Transferase.
 SQ SEQUENCE 1133 AA; 129323 MW; 32B72C4E429B20B9 CRC64;

Query Match 77.3%; Score 34; DB 10; Length 1133;
 Best Local Similarity 66.7%; Pred. No. 2.9e+02;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 HVSAALGNL 9
 ||:|||||
 Db 656 HLSILGNL 664

RESULT 11

Q9HVP1 PRELIMINARY; PRT; 149 AA.

ID Q9HVP1; AC Q9HVP1; DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Hypothetical protein PA4536.
 GN PA4536.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 OX NCBI_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 15692 / PA01;
 RX MEDLINE=20437337; PubMed=10984043;
 RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
 RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
 RA Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.P.,
 RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
 RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
 RT opportunistic pathogen."
 RL Nature 406:959-964(2000).
 DR EMBL; AE004867; AAG07924.1; -
 DR InterPro; IPR005302; MOSC.
 DR Pfam; PF03473; MOSC; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 149 AA; 16542 MW; EA5952E315E0946D CRC64;

Query Match 75.0%; Score 33; DB 16; Length 149;
 Best Local Similarity 66.7%; Pred. No. 53;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 HVSAALGNL 9
 ||:|||||
 Db 51 HIEALGTL 59

RESULT 12

Q8ITF4 PRELIMINARY; PRT; 255 AA.

ID Q8ITF4; AC Q8ITF4; DT 01-MAR-2003 (TREMBLrel. 23, Created)
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE SALP-1.
 OS Giardia lamblia (Giardia intestinalis).
 OC Eukaryota; Diplomonadida; Hexamitidae; Giardiinae; Giardia.
 OX NCBI_TaxID=5741;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Svard S.G., Weiland M.E.L., Palm D.J.E.;
 RT "Immunoreactive proteins during acute human giardiasis."
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF514363; AAN73280.1; -
 SQ SEQUENCE 255 AA; 29827 MW; 8AAD9E8A66405BD8 CRC64;

Query Match 75.0%; Score 33; DB 5; Length 255;
 Best Local Similarity 66.7%; Pred. No. 95;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 HVSAALGNL 9
 ||:|||||
 Db 81 HVTALNNLI 89

RESULT 13

Q98315 PRELIMINARY; PRT; 308 AA.

ID Q98315

```
AC Q98315; O12898; O12617;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DE MC149R.
DE 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
GN MC149R OR H-M-N-4.
OS Molluscum contagiosum virus subtype 1 (MCV1).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Molluscipoxvirus.
OX NCBI_TaxID=10280;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96325459; PubMed=8670425;
RA Senkevich T.G., Bugert J.J., Sisler J.R., Koonin E.V., Darai G.,
RA Moss B.;
RT "Genome sequence of a human tumorigenic poxvirus: prediction of
RT specific host response-evasion genes.";
RT Science 273:813-816(1996).
RN [2]
RP SEQUENCE FROM N.A.
RA Senkevich T.G., Bugert J.J., Sisler J.R., Koonin E.V., Darai G.,
RA Moss B.;
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Moratilla M., Agromayor M., Nunez A., Funes J.M., Varas A.J.,
RA Lopez-Esteban J.L., Esteban M., Martin-Gallardo A.;
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U60315; AAC55277.1; -.
DR EMBL; U86945; AAB58020.1; -.
SQ SEQUENCE 308 AA; 33979 MW; B7943721F09CB613 CRC64;

Query Match 75.0%; Score 33; DB 12; Length 308;
Best Local Similarity 77.8%; Pred. No. 1.2e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 HVSAAGNLL 9
DB 174 HVSAAGDGLL 182

RESULT 14
Q8CHX3 PRELIMINARY; PRT; 559 AA.
AC Q8CHX3;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE RIKEN cDNA 6720467009 gene.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CZECH II;
RA Strausberg R.;
RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC038386; AAB38386.1; -.
SQ SEQUENCE 559 AA; 64441 MW; D2FEF972511CFC40 CRC64;

Query Match 75.0%; Score 33; DB 11; Length 559;
Best Local Similarity 75.0%; Pred. No. 2.2e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 HVSAAGNLL 8
DB 551 HVSKMGNLL 558

RESULT 15
Q8PXC9 PRELIMINARY; PRT; 623 AA.
ID Q8PXC9
```

```
AC Q8PXC9;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DE 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Archaeosine tRNA-ribosyltransferase (EC 2.4.2.-).
GN MML292.
OS Methanosarcina mazel (Methanosarcina frisia).
OC Archaea; Euryarchaeota; Methanococci; Methanosarcinales;
OC Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxID=2209;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Goel / Go1 / ATCC BAA-199 / DSM 3647 / OCM 88;
RX MEDLINE=22120827; PubMed=12125824;
RA Deppenmeier U., Johann A., Hartsch T., Merkl R., Schmitz R.A.,
RA Martinez-Arias R., Henne A., Wierzer A., Baeumer S., Jacobi C.,
RA Brueggemann H., Lienard T., Christmann A., Boemecke M., Steckel S.,
RA Bhattacharyya A., Lykidis A., Overbeek R., Klenk H.-P., Gunsalus R.P.,
RA Fritz H.-J., Gottschalk G.;
RT "The genome of Methanosarcina mazel: evidence for lateral gene
RT transfer between Bacteria and Archaea.";
RL J. Mol. Microbiol. Biotechnol. 4:453-461(2002).
DR EMBL; AE013360; AAM30988.1; -.
DR InterPro; IPR002478; PUA.
DR InterPro; IPR002616; tRNA_ribo_trans.
DR Pfam; PF01472; PUA; 1.
DR Pfam; PF01702; TGT; 1.
DR TIGRFAMs; TIGR00449; tgt_general; 1.
DR TIGRFAMs; TIGR00451; unchar_dom_2; 1.
DR TIGRFAMs; TIGR00451; unchar_dom_2; 1.
KW Transferase; Glycosyltransferase; Complete proteome.
SQ SEQUENCE 623 AA; 69096 MW; C25881CFF210B586 CRC64;

Query Match 75.0%; Score 33; DB 17; Length 623;
Best Local Similarity 75.0%; Pred. No. 2.5e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 HVSAAGNLL 8
DB 32 HTAALGNL 39

Search completed: January 20, 2004, 17:01:03
Job time : 27.9695 secs
```

